Comptes rendus mathématiques
Mathematical Reports

JUNE / JUIN 2000

IN THIS ISSUE / DANS CE NUMÉRO

22
No 2

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 33–54

THE ORBIT BUNDLE PICTURE OF COTANGENT BUNDLE REDUCTION

JERROLD E. MARSDEN, FRSC, AND MATTHEW PERLMUTTER

ABSTRACT. Cotangent bundle reduction theory is a basic and well developed subject in which one performs symplectic reduction on cotangent bundles. One starts with a (free and proper) action of a Lie group G on a configuration manifold Q, considers its natural cotangent lift to T^*Q and then one seeks realizations of the corresponding symplectic or Poisson reduced space. We further develop this theory by explicitly identifying the symplectic leaves of the Poisson manifold T^*Q/G, decomposed as a Whitney sum bundle, $T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$ over Q/G. The splitting arises naturally from a choice of connection on the G-principal bundle $Q \to Q/G$. The symplectic leaves are computed and a formula for the reduced symplectic form is found.

1. Introduction and overview.

A BRIEF HISTORY OF REDUCTION THEORY. Reduction theory for mechanical systems with symmetry has its origins in the classical work of Euler and Lagrange in the late 1700s and that of Hamilton, Jacobi, Routh and Poincaré in the period 1830–1910. The immediate goal of reduction theory is to use conservation laws and the associated symmetries to reduce the number of dimensions required to describe a mechanical system. For example, using rotational invariance and conservation of angular momentum, the classical central force problem for a particle moving in $\mathbb{R}^3$ (a three degree of freedom mechanical system) can be reduced to a single second order ordinary differential equation in the radial variable, describing a new, *reduced* mechanical system with just one degree of freedom.

By 1830, variational principles, such as Hamilton's principle and the principle of least action, as well as canonical Poisson brackets were fairly well understood and there were shades of symplectic geometry already in the work of Lagrange. Several classical examples of reduction were understood in that era, such as the elimination of cyclic variables, which we would call today reduction by Abelian groups, which was primarily due to Routh, as well as Jacobi's elimination of the node for interacting particles in $\mathbb{R}^3$ with SO(3) symmetry.

Lie, by 1890, deepened the mathematical understanding of symplectic and Poisson geometry and their link with symmetry. Between 1901 and 1910, Poincaré discovered how to generalize the Euler equations for rigid body mechanics and fluids to general Lie algebras.

Received by the editors May 3, 1999.
AMS subject classification: 37J15, 53D20, 70H33.

Interestingly, these methods, with some exceptions, remained nearly dormant since the time of Poincaré, for over half a century. Meanwhile, Cartan and others developed the needed tools of differential forms and analysis on manifolds, setting the stage for the modern era of reduction theory. Naturally, much attention was also going to exciting developments in relativity theory and quantum mechanics. This modern era began with the fundamental papers of Arnold [4] and Smale [73]. Arnold focussed on systems on Lie algebras and their duals, as in the works of Lie and Poincaré, while Smale focussed on the Abelian case giving, in effect, a modern version of Routh reduction.

With hindsight, we now know that the description of many physical systems such as rigid bodies and fluids requires *noncanonical Poisson brackets* and *constrained variational principles* of the sort studied by Lie and Poincaré. One example of noncanonical Poisson brackets on $\mathfrak{g}^*$, the dual of a Lie algebra $\mathfrak{g}$, are called, following Marsden and Weinstein [47], *Lie-Poisson brackets*. These structures were known to Lie around 1890, although Lie apparently did not recognize their importance in mechanics. The symplectic leaves in these structures, namely the coadjoint orbit symplectic structures, although implicit in Lie's work, were discovered by Kirillov, Kostant, and Souriau in the 1960's.

To synthesize the Lie algebra reduction methods of Arnold [4] with the techniques of Smale [73] on the reduction of cotangent bundles by Abelian groups, Marsden and Weinstein [45] developed reduction theory in the general context of symplectic manifolds and equivariant momentum maps; related results, but with a different motivation and construction (not stressing equivariant momentum maps) were found by Meyer [49].

THE SYMPLECTIC REDUCED SPACE. The construction of the symplectic reduced space is now standard: let (P, Ω) be a symplectic manifold and let a Lie group G act freely and properly on P by symplectic maps. The free and proper assumption is needed if one wishes to avoid singularities in the reduction procedure. Assume that this action has an equivariant momentum map $\mathbf{J}: P \to \mathfrak{g}^*$. Then the *symplectic reduced space* $\mathbf{J}^{-1}(\mu)/G_\mu = P_\mu$ is a symplectic manifold in a natural way; the induced symplectic form Ω_μ is determined uniquely by $\pi_\mu^*\Omega_\mu = i_\mu^*\Omega$ where $\pi_\mu: \mathbf{J}^{-1}(\mu) \to P_\mu$ is the projection and $i_\mu: \mathbf{J}^{-1}(\mu) \to P$ is the inclusion. If the momentum map is not equivariant, Souriau [74] discovered how to centrally extend the group (or algebra) to make it equivariant.

Coadjoint orbits were shown to be symplectic reduced spaces by Marsden and Weinstein [45]: in the reduction construction, one chooses $P = T^*G$, with G acting by (say left) translation and the corresponding space P_μ is identified with the coadjoint orbit $\mathcal{O}_\mu$ through μ together with its coadjoint orbit symplectic structure. Likewise, the Lie-Poisson bracket on $\mathfrak{g}^*$ is inherited from the canonical Poisson structure on T^*G by *Poisson reduction*, that is, by simply identifying $\mathfrak{g}^*$ with the quotient $(T^*G)/G$. It is not clear who first *explicitly* observed this, but it is *implicit* in many works such as Lie [31], Kirillov [26], [27], Guillemin and

Sternberg [18], and Marsden and Weinstein [46], [47], but is *explicit* in Marsden, Weinstein, Ratiu, Schmid, and Spencer [48] and in Holmes and Marsden [22].

Kazhdan, Kostant and Sternberg [25] showed that P_μ is symplectically diffeomorphic to an orbit reduced space $P_\mu \cong J^{-1}(\mathcal{O}_\mu)/G$ and from this it follows that P_μ are the symplectic leaves in P/G. This paper was also one of the first to notice deep links between reduction and integrable systems, a subject continued by, for example, Bobenko, Reyman and Semenov-Tian-Shansky [10] in their spectacular group theoretic explanation of the integrability of the Kowalewski top.

The way in which the *Poisson* structure on P_μ is related to that on P/G was clarified in a generalization of *Poisson reduction* due to Marsden and Ratiu [39], a technique that has also proven useful in integrable systems (see, *e.g.*, Pedroni [59] and Vanhaecke [76]).

LAGRANGIAN REDUCTION. Routh reduction for Lagrangian systems is classically associated with systems having cyclic variables (this is almost synonymous with having an Abelian symmetry group); modern accounts can be found in Arnold [5] and in Marsden and Ratiu [38, §8.9]. A key feature of Routh reduction is that when one drops the Euler-Lagrange equations to the quotient space associated with the symmetry, and when the momentum map is constrained to a specified value (*i.e.*, when the cyclic variables and their velocities are eliminated using the given value of the momentum), then the resulting equations are in Euler-Lagrange form *not* with respect to the Lagrangian itself, but with respect to the *Routhian*. In his classical work, Routh [68] applied these ideas to stability theory, a precursor to the energy-momentum method for stability (Simo, Lewis, and Marsden [71]; see Marsden [33] for an exposition and references). Of course, Routh's stability method is still widely used in mechanics.

Another key ingredient in Lagrangian reduction is the classical work of Poincaré [60] in which the *Euler-Poincaré equations* were introduced. Poincaré realized that both the equations of fluid mechanics and the rigid body and heavy top equations could all be described in Lie algebraic terms in a beautiful way.

TANGENT AND COTANGENT BUNDLE REDUCTION. The simplest case of cotangent bundle reduction is reduction at zero in which case one chooses $P = T^*Q$ and then the symplectic reduced space formed at $\mu = 0$ is given by $P_0 = T^*(Q/G)$, the latter with the canonical symplectic form. Another basic case is when G is Abelian. Here, $(T^*Q)_\mu \cong T^*(Q/G)$ but the latter has a symplectic structure modified by magnetic terms; that is, by the curvature of the mechanical connection.

The Abelian version of cotangent bundle reduction was developed by Smale [73] and Satzer [70] and was generalized to the nonabelian case in Abraham and Marsden [1]. Kummer [29] introduced the interpretations of these results in terms of a connection, now called the *mechanical connection*. The geometry of this situation was used to great effect in, for example, Guichardet [17], Iwai [23] and Montgomery [50], [53], [54].

Tangent and cotangent bundle reduction evolved into what we now term as the "bundle picture" or the "gauge theory of mechanics". This picture was first developed by Montgomery, Marsden and Ratiu [57] and Montgomery [50], [51]. That work was motivated and influenced by the work of Sternberg [75] and Weinstein [77] on a *Yang-Mills construction* that is in turn motivated by Wong's equations, that is, the equations for a particle moving in a Yang-Mills field. The main result of the bundle picture gives a structure to the quotient spaces $(T^*Q)/G$ and $(TQ)/G$ when G acts by the cotangent and tangent lifted actions.

SEMIDIRECT PRODUCT REDUCTION. In the simplest case of a semidirect product, one has a Lie group G that acts on a vector space V (and hence on its dual V^*) and then one forms the semidirect product $S = G \circledS V$, generalizing the semidirect product structure of the Euclidean group $\mathrm{SE}(3) = \mathrm{SO}(3) \circledS \mathbb{R}^3$.

Consider the isotropy group G_{a_0} for some $a_0 \in V^*$. The *semidirect product reduction theorem* states that each of the *symplectic reduced spaces for the action of G_{a_0} on T^*G is symplectically diffeomorphic to a coadjoint orbit in $(\mathfrak{g} \circledS V)^*$, the dual of the Lie algebra of the semi-direct product.* This semidirect product theory was developed by Guillemin and Sternberg [18], Ratiu [62], [65], [66], and Marsden, Ratiu and Weinstein [41], [42].

This construction is used in applications where one has "advected quantities" (such as the direction of gravity in the heavy top, density in compressible flow and the magnetic field in MHD). Its Lagrangian counterpart was developed in Holm, Marsden, and Ratiu [20] along with applications to continuum mechanics. Cendra, Holm, Hoyle, and Marsden [11] applied this idea to the Maxwell-Vlasov equations of plasma physics. Cendra, Holm, Marsden, and Ratiu [12] showed how Lagrangian semidirect product theory it fits into the general framework of Lagrangian reduction.

REDUCTION BY STAGES AND GROUP EXTENSIONS. The semidirect product reduction theorem can be viewed using reduction by stages: one reduces T^*S by the action of the semidirect product group $S = G \circledS V$ in two stages, first by the action of V at a point a_0 and followed by the action of G_{a_0}. Semidirect product reduction by stages for actions of semidirect products on general symplectic manifolds was developed and applied to underwater vehicle dynamics in Leonard and Marsden [30]. Motivated partly by semidirect product reduction, Marsden, Misiolek, Perlmutter, and Ratiu [35], [36] gave a significant generalization of semidirect product theory in which one has a group M with a normal subgroup $N \subset M$ (so M is a group extension of N) and M acts on a symplectic manifold P. One wants to reduce P in two stages, first by N and then by M/N. On the Poisson level this is easy: $P/M \cong (P/N)/(M/N)$ but on the symplectic level it is quite subtle.

Cotangent bundle reduction by stages is especially interesting for group extensions. An example of such a group, besides semidirect products, is the Bott-

Virasoro group, where the Gelfand-Fuchs cocycle may be interpreted as the curvature of a mechanical connection.

LAGRANGE-POINCARÉ AND LAGRANGE-ROUTH REDUCTION. Marsden and Scheurle [43], [44] showed how to generalize the Routh theory to the nonabelian case as well as showing how to get the Euler-Poincaré equations for matrix groups by the important technique of *reducing variational principles.* This approach was motivated by Cendra and Marsden [14] and Cendra, Ibort, and Marsden [13]. The work of Bloch, Krishnaprasad, Marsden, and Ratiu [9] generalized the Euler-Poincaré variational structure to general Lie groups and Cendra, Marsden, and Ratiu [15] carried out a Lagrangian reduction theory that extends the Euler-Poincaré case to arbitrary configuration manifolds. This work was in the context of the Lagrangian analogue of Poisson reduction in the sense that no momentum map constraint is imposed.

One of the things that makes the Lagrangian side of the reduction story interesting is the lack of a general category that is the Lagrangian analogue of Poisson manifolds. Such a category, that of *Lagrange-Poincaré bundles* is developed in Cendra, Marsden, and Ratiu [15], with the tangent bundle of a configuration manifold and a Lie algebra as its most basic examples. That work also develops the Lagrangian analogue of reduction for central extensions and, as in the case of symplectic reduction by stages mentioned above, cocycles and curvatures enter in this context in a natural way.

The Lagrangian analogue of the bundle picture is the bundle $(TQ)/G$, a vector bundle over Q/G; this bundle was studied in Cendra, Marsden, and Ratiu [15]. In particular, the equations and variational principles are developed on this space. For $Q = G$ this reduces to Euler-Poincaré reduction. A G-invariant Lagrangian L on TQ induces a Lagrangian l on $(TQ)/G$. The resulting equations inherited on this space, given explicitly later, are the *Lagrange-Poincaré equations* (or the *reduced Euler-Lagrange equations*).

The above results develop what we might call *Lagrange-Poincaré reduction.* On the other hand, in the nonabelian Routh reduction theory of Marsden and Scheurle [43] one imposes, as in symplectic reduction, a momentum map constraint. This was put into the bundle context by Jalnapurkar and Marsden [24] and Marsden, Ratiu and Scheurle [40]. We may call this *Lagrange-Routh reduction.*

APPLICATIONS OF REDUCTION THEORY. Reduction theory for mechanical systems with symmetry has proven to be a powerful tool enabling advances in stability theory (from the Arnold method to the energy-momentum method) as well as in bifurcation theory of mechanical systems, geometric phases via reconstruction—the inverse of reduction—as well as uses in control theory from stabilization results to a deeper understanding of locomotion. Methods of Lagrangian reduction have proven very useful in, for example, optimal control problems. It was used in Koon and Marsden [28] to extend the falling cat theorem of

Montgomery [53] to the case of nonholonomic systems. For a general introduction to some of these ideas and for further references, see, for example, Marsden and Ratiu [38], Leonard and Marsden [30] and Marsden and Ostrowski [37].

SINGULAR REDUCTION. Singular reduction starts with the observation of Smale [73] that $z \in P$ is a regular point of $\mathbf{J}$ iff z has no continuous isotropy. Motivated by this, Arms, Marsden, and Moncrief [2] showed that the level sets $\mathbf{J}^{-1}(0)$ of an equivariant momentum map $\mathbf{J}$ have quadratic singularities at points with continuous symmetry. While such a result is easy for compact group actions on finite dimensional manifolds, the main examples of Arms, Marsden, and Moncrief [2], [3] were *infinite dimensional*—both the phase space and the group. We will not be concerned with singular reduction in this paper; we refer to Ortega and Ratiu [58] for further references and discussion.

There are of course many other aspects of reduction theory and associated techniques that we do not attempt to review here, including resonant systems, nonholonomic mechanics, the method of invariants, *etc.*

THE MAIN RESULT. The main new result of this paper is Theorem 4.3. This gives an expression for the reduced symplectic form on the symplectic leaves of $(T^*Q)/G$, each of which is determined to be a fiber products of the form $T^*(Q/G) \times_{Q/G} \tilde{\mathcal{O}}$, where $\tilde{\mathcal{O}}$ is the associated coadjoint orbit bundle. The symplectic structure restricted to the orbit bundle involves the curvature of the connection, the orbit symplectic form, and interaction terms that pair tangent vectors to the orbit with the vertical projections of tangent vectors to the configuration space. Our result may be viewed as a symplectic version of the global Poisson bracket formula on reduced cotangent bundles due to Montgomery, Marsden and Ratiu [57] and Montgomery [51]; see also Blaom [8] and Zaalani [79] for related results.

2. The symplectic leaves.

Throughout the paper, we let a Lie group G act freely and properly on a manifold Q so that the natural quotient map $\pi: Q \to Q/G$ defines a principal bundle. Let $\mathcal{A}$ be a principal connection this bundle and let $\tilde{\mathfrak{g}}$ denote the associated bundle to the Lie algebra $\mathfrak{g}$, namely $\tilde{\mathfrak{g}} = (Q \times \mathfrak{g})/G$, which we regard as a vector bundle over Q/G. We recall the following natural bundle isomorphisms (see Cendra, Holm, Marsden and Ratiu [12]):

LEMMA 2.1. *There are bundle isomorphisms*

$$\alpha_{\mathcal{A}}: TQ/G \to T(Q/G) \oplus \tilde{\mathfrak{g}} \quad \textit{and} \quad (\alpha_{\mathcal{A}}^{-1})^*: T^*Q/G \to T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*. \tag{2.1}$$

PROOF. Given $v_q \in T_qQ$, denote its equivalence class in $T(Q/G)$ by $[v_q]$. We claim that the mapping $\alpha_{\mathcal{A}}: [v_q] \mapsto T_q\pi(v_q) \oplus [q, \mathcal{A}(q)(v_q)]$ is well defined and induces the desired isomorphism of TQ/G with $T(Q/G) \oplus \tilde{\mathfrak{g}}$. To see this, and to help clarify notations in the sequel, consider another representative of the orbit

$[v_q]$, given by $g \cdot v_q$ where we use concatenated notation for the tangent lifted action of G on TQ. We have $T_{g\cdot q}\pi(g \cdot v_q) = T_q\pi(v_q)$ and

$$\begin{aligned}[g \cdot q, \mathcal{A}(g \cdot q)(g \cdot v_q)] &= [g \cdot q, (\phi_g^* \mathcal{A})] = [g \cdot q, \mathrm{Ad}_g \mathcal{A}(q)(v_q)] \\ &= [q, \mathcal{A}(q)(v_q)].\end{aligned}$$

The inverse of this map is given by $v_{[q]} \oplus [q, \xi] \mapsto [\mathrm{hor}_q v_{[q]} + \xi_Q(q)]$ as is readily verified. We therefore have a vector bundle isomorphism.

We next compute $(\alpha_{\mathcal{A}}^{-1})^*: T^*Q/G \to T^*(Q/G) \oplus \tilde{\mathfrak{g}}^*$, the dual of the inverse map:

$$\begin{aligned}\left\langle (\alpha_{\mathcal{A}}^{-1})^* \left([\alpha_q], (u_{[q]}, [q, \xi])\right) \right\rangle &= \langle [\alpha_q], [\mathrm{hor}_q \cdot u_{[q]} + \xi_Q(q)] \rangle \\ &= \langle \alpha_q, \mathrm{hor}_q \cdot u_{[q]} \rangle + \langle \alpha_q, \xi_Q(q) \rangle \\ &= \langle \mathrm{hor}_q^* \alpha_q, u_{[q]} \rangle + \langle \mathbf{J}(\alpha_q), \xi \rangle,\end{aligned}$$

where $\mathrm{hor}_q^*: T_q^*Q \to T_{[q]}^*(Q/G)$ is dual to the horizontal lift map $\mathrm{hor}_q: T_{[q]}(Q/G) \to T_qQ$ so that we conclude $(\alpha_{\mathcal{A}}^{-1})^*([\alpha_q]) = \left(\mathrm{hor}_q^* \alpha_q, [q, \mathbf{J}(\alpha_q)]\right)$. ∎

This bundle isomorphism can be recast as follows (see Cushman and Śniatycki [16]). Consider the maps

$$\Delta: T^*Q \to T^*Q/G \to \tilde{\mathfrak{g}}^*; \quad \alpha_q \mapsto [\alpha_q] \mapsto [q, \mathbf{J}(\alpha_q)]$$

and

$$\Gamma: T^*Q \to T^*Q/G \to T^*(Q/G); \quad \alpha_q \mapsto [\alpha_q] \mapsto \mathrm{hor}_q^* \alpha_q.$$

Notice that the map $\alpha_q \mapsto \mathrm{hor}_q^* \alpha_q$ drops to T^*Q/G, since we have, for all $V_{[q]} \in T_{[q]}(Q/G)$,

$$\begin{aligned}\langle \mathrm{hor}_{g\cdot q}^*(g \cdot \alpha_q), V_{[q]} \rangle &= \langle g \cdot \alpha_q, \mathrm{hor}_{g\cdot q} V_{[q]} \rangle = \langle \alpha_q, g^{-1} \cdot \mathrm{hor}_{g\cdot q} V_{[q]} \rangle \\ &= \langle \alpha_q, g^{-1} \cdot (g \cdot \mathrm{hor}_q V_{[q]} \rangle = \langle \alpha_q, \mathrm{hor}_q V_{[q]} \rangle\end{aligned}$$

where we use the fact that $g \cdot \mathrm{hor}_q = \mathrm{hor}_{g\cdot q}$.

We can write $(\alpha_{\mathcal{A}}^{-1})^* = \Gamma \oplus \Delta$. A partial inverse to the projection Δ is given in the next proposition.

PROPOSITION 2.2. *Consider the map,*

$$\sigma: Q \times \mathfrak{g}^* \to T^*Q/G \tag{2.2}$$

given by $(q, \nu) \mapsto \mathcal{A}(q)^* \nu$. This map is equivariant with respect to the diagonal action of G on $Q \times \mathfrak{g}^*$ and the cotangent lifted action of G on T^*Q, and so uniquely defines a map on the quotient,

$$\tilde{\sigma}: \tilde{\mathfrak{g}}^* \to T^*Q/G. \tag{2.3}$$

This is a fiber preserving bundle map which is injective on each fiber and satisfies $\Delta \circ \tilde{\sigma} = \mathrm{id}\,|_{\tilde{\mathfrak{g}}^*}$.

PROOF. Under the map σ, $g \cdot (q, \nu) = (g \cdot q, \mathrm{Ad}^*_{g^{-1}} \nu) \mapsto \mathcal{A}(g \cdot q)^*(\mathrm{Ad}^*_{g^{-1}} \nu)$. However, for all $v \in T_{g \cdot q}Q$,

$$\begin{aligned}\langle \mathcal{A}(g \cdot q)^*(\mathrm{Ad}^*_{g^{-1}} \nu), v\rangle &= \langle \mathrm{Ad}^*_{g^{-1}} \nu, \mathcal{A}(g \cdot q)v\rangle = \langle \nu, \mathrm{Ad}_{g^{-1}} \mathcal{A}(g \cdot q)v\rangle \\ &= \langle \nu, (\psi_{g^{-1}})^* \mathcal{A}(g \cdot q)v\rangle = \langle \nu, \mathcal{A}(q)(T\psi_{g^{-1}}v)\rangle \\ &= \langle g \cdot \mathcal{A}(q)^*\nu, v\rangle,\end{aligned}$$

from which we conclude equivariance of σ. Also, for each $[q, \nu] \in \tilde{\mathfrak{g}}^*$,

$$\Delta(\tilde{\sigma}([q, \nu])) = \Delta([\mathcal{A}(q)^*\nu]) = [q, \mathbf{J}(\mathcal{A}(q)^*\nu)] = [q, \nu],$$

since, for all $\xi \in \mathfrak{g}$, $\langle \mathbf{J}(\mathcal{A}(q)^*\nu), \xi\rangle = \langle \nu, \mathcal{A}(q)(\xi_Q(q))\rangle = \langle \nu, \xi\rangle$. ∎

We next determine the image, under $(\alpha_{\mathcal{A}}^{-1})^*$ of the symplectic leaves of the Poisson manifold T^*Q/G, which we know from the symplectic correspondence theorem (see Weinstein [78], Blaom [7]) are given by $\mathbf{J}^{-1}(\mathcal{O})/G$ for each coadjoint orbit $\mathcal{O}$ in $\mathfrak{g}^*$.

THEOREM 2.3. *We have* $(\alpha_{\mathcal{A}}^{-1})^*(\mathbf{J}^{-1}(\mathcal{O})/G) = T^*(Q/G) \times_{Q/G} \tilde{\mathcal{O}}$, *where* $\tilde{\mathcal{O}} = (Q \times \mathcal{O})/G$ *is the associated bundle using the coadjoint action of* G *on* $\mathcal{O}$.

PROOF. From the definition of the bundle isomorphism $\alpha_{\mathcal{A}}$,

$$\begin{aligned}(\alpha_{\mathcal{A}}^{-1})^*(\mathbf{J}^{-1}(\mathcal{O})/G) &= \{(\Gamma(\alpha_q), \Delta(\alpha_q)) \mid \mathbf{J}(\alpha_q) \in \mathcal{O}\} \\ &= \{(\mathrm{hor}_q^* \alpha_q, [q, \mathbf{J}(\alpha_q)]) \mid \mathbf{J}(\alpha_q) \in \mathcal{O}\}.\end{aligned}$$

First we characterize the sets $T_q^*Q \cap \mathbf{J}^{-1}(\mathcal{O})$, using the connection $\mathcal{A}$. Denote by $\mathbf{J}_q$, the restriction of $\mathbf{J}$ to T_q^*Q, and let $\sigma_q \colon \mathfrak{g} \mapsto T_qQ$, be the injective infinitesimal generator map. Using the fact that $\mathbf{J}_q = \sigma_q^*$, we have for all $\xi \in \mathfrak{g}$

$$\langle \mathbf{J}(\alpha_q + \mathcal{A}_\mu(q)), \xi\rangle = \langle \alpha_q, \xi_Q(q)\rangle + \langle \mu, \xi\rangle = \langle \mu, \xi\rangle \tag{2.4}$$

where the second equality holds since $\alpha_q \in V^0$, the annihilator of the vertical sub-bundle of TQ. We conclude that

$$T_q^*Q \cap \mathbf{J}^{-1}(\mathcal{O}) = \{V_q^0 + \mathcal{A}_\mu(q) \mid \mu \in \mathcal{O}\}.$$

Recall that since σ_q^* is surjective, $\tau \circ \mathbf{J}^{-1}(\mathcal{O}) = Q$, where $\tau_{T^*Q} \colon T^*Q \to Q$ is the cotangent bundle projection. Now apply hor_q^* to each fiber over Q in $\mathbf{J}^{-1}(\mathcal{O})$. That is, for each $q \in Q$, we consider

$$\mathrm{hor}_q^*(\mathbf{J}^{-1}(\mathcal{O}) \cap T_q^*Q). \tag{2.5}$$

First, note that for all $X_{[q]} \in T_{[q]}(Q/G)$,

$$\langle \mathrm{hor}_q^*(\mathcal{A}_\mu(q)), X_{[q]}\rangle = \langle \mu, \mathcal{A}(q)(\mathrm{hor}_q(X_{[q]}))\rangle = 0 \tag{2.6}$$

so that $\mathrm{hor}_q^*(\mathcal{A}_\mu(q)) = 0$. Furthermore, since hor_q is injective, $\mathrm{hor}_q^*: T_q^*Q \to T_{[q]}^*(Q/G)$ is surjective with $\ker \mathrm{hor}_q^* = H^0$, where H^0 denotes the annihilator of the horizontal subbundle of TQ. Thus, as a linear map, $\mathrm{hor}_q^*: V^0 \to T_{[q]}^*(Q/G)$ is an isomorphism. Consider the set of pairs, $\{(\Gamma(\alpha_q), \Delta(\alpha_q)) \mid \mathbf{J}(\alpha_q) \in \mathcal{O}\}$. Each α_q can be uniquely expressed as $\beta_q + \mathcal{A}_\mu(q)$ for some $\beta_q \in V^0$ and $\mu \in \mathcal{O}$. For a fixed μ, let β_q range over V_q^0. This generates the set $T_{[q]}^*(Q/G) \times [q, \mu]$ since $\mathbf{J}$ vanishes on V^0. The result now follows by varying $\mu \in \mathcal{O}$. ∎

3. **Orbit reduction.** Let us recall the characterizing property of the reduced symplectic forms in the orbit reduction setting of Kazhdan, Kostant and Sternberg [25].

THEOREM 3.1. *Let μ be a regular value of an equivariant momentum map $\mathbf{J}: P \to \mathfrak{g}^*$ of a left symplectic action of G on the symplectic manifold (P, Ω) and assume that the symplectic reduced space P_μ is a manifold with π_μ a submersion. Let $\mathcal{O}$ be the coadjoint orbit through μ in $\mathfrak{g}_+^*$. Then*

1. *$\mathbf{J}$ is transversal to $\mathcal{O}$ so $\mathbf{J}^{-1}(\mathcal{O})$ is a manifold;*
2. *$\mathbf{J}^{-1}(\mathcal{O})/G$ has a unique differentiable structure such that the canonical projection $\pi_{\mathcal{O}}: \mathbf{J}^{-1}(\mathcal{O}) \to \mathbf{J}^{-1}(\mathcal{O})/G$ is a surjective submersion;*
3. *there is a unique symplectic structure $\Omega_{\mathcal{O}}$ on $\mathbf{J}^{-1}(\mathcal{O})/G$ such that*

$$\iota_{\mathcal{O}}^*\Omega = \pi_{\mathcal{O}}^*\Omega_{\mathcal{O}} + \mathbf{J}_{\mathcal{O}}^*\omega_{\mathcal{O}}^+, \tag{3.1}$$

where $\iota_{\mathcal{O}}: \mathbf{J}^{-1}(\mathcal{O}) \to P$ is the inclusion, $\mathbf{J}_{\mathcal{O}} = \mathbf{J}|\mathbf{J}^{-1}(\mathcal{O})$, and $\omega_{\mathcal{O}}^+$ is the "+" orbit symplectic structure on $\mathcal{O}$.

By considering a momentum shift we can realize a bundle isomorphism between $\mathbf{J}^{-1}(\mathcal{O})$ and the space $V^0 \times \mathcal{O}$. Since it will be shown that this isomorphism is G equivariant, it determines a unique diffeomorphism between $V^0 \times \mathcal{O}/G$ and $\mathbf{J}^{-1}(\mathcal{O})/G$. We will characterize the symplectic form on the former by pulling back the "characterizing" symplectic form on $\mathbf{J}^{-1}(\mathcal{O})$. Furthermore, it will be shown that $V^0 \times \mathcal{O}/G$ is diffeomorphic to $T^*(Q/G) \oplus \tilde{\mathcal{O}}$, so that the reduced symplectic form can be expressed on this space as well. Since $\mathbf{J}^{-1}(\mathcal{O}) \subset T^*Q$, the reduced symplectic form is determined by the restriction of the canonical symplectic form in T^*Q to $\mathbf{J}^{-1}(\mathcal{O})$, which in turn is determined by the restriction of the canonical one-form to $\mathbf{J}^{-1}(\mathcal{O})$.

LEMMA 3.2. *There is a G-equivariant bundle isomorphism,*

$$\chi: V^0 \times \mathcal{O} \to \mathbf{J}^{-1}(\mathcal{O}) \tag{3.2}$$

that uniquely determines a diffeomorphism, $\bar{\chi}$ on the quotient spaces so that the following diagram commutes

$$\begin{array}{ccc} V^0 \times \mathcal{O} & \xrightarrow{\chi} & \mathbf{J}^{-1}(\mathcal{O}) \\ {\scriptstyle \tilde{\pi}_G}\downarrow & & \downarrow{\scriptstyle \pi_{\mathcal{O}}} \\ (V^0 \times \mathcal{O})/G & \xrightarrow[\bar{\chi}]{} & \mathbf{J}^{-1}(\mathcal{O})/G \end{array} .$$

PROOF. The map χ is given by

$$\chi(\alpha_q, \nu) = \alpha_q + \mathcal{A}(q)^* \nu. \tag{3.3}$$

This map takes values in $\mathbf{J}^{-1}(\mathcal{O})$, as is seen from the proof of the previous theorem. From the characterization of the fibers of the bundle $\mathbf{J}^{-1}(\mathcal{O}) \to Q$, it follows that this map is onto. Since it is a momentum shift, it is clearly invertible with inverse

$$\alpha_q \mapsto \alpha_q - \mathcal{A}(q)^* \mathbf{J}(\alpha_q). \tag{3.4}$$

We check G-equivariance as follows:

$$\begin{aligned} \chi\big(g \cdot (\alpha_q, \nu)\big) &= \chi(g \cdot \alpha_q, g \cdot \nu) = g \cdot \alpha_q + \mathcal{A}(g \cdot q)^* (g \cdot \nu) \\ &= g \cdot \alpha_q + g \cdot \big(\mathcal{A}(q)^* \nu\big) = g \cdot \big(\chi(\alpha_q, \nu)\big) \end{aligned}$$

where the third equality uses the invariance properties of the connection. ∎

Because χ is G-equivariant, the pull back by χ of the G-invariant form on $\mathbf{J}^{-1}(\mathcal{O})$, $\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}$, given by $\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}})$, is a G-invariant form on $V^0 \times \mathcal{O}$. In fact, the form drops to the quotient by the diagonal G action since

$$\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}) = \bar{\pi}_G^* \bar{\chi}^* \Omega_{\mathcal{O}}, \tag{3.5}$$

where $\bar{\pi}_G \colon V^0 \times \mathcal{O} \to V^0 \times \mathcal{O}/G$ denotes the projection. This follows since the diagram in the preceding theorem is commutative.

3.1. *The two-form on* $V^0 \times \mathcal{O}$. We proceed to characterize the structure of the two-form $\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}})$ on $V^0 \times \mathcal{O}$. By construction,

$$\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}) = \chi^* \iota_{\mathcal{O}}^* \Omega - \chi^* \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+. \tag{3.6}$$

THE SECOND TERM. We claim that

$$\chi^* \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+ = \pi_2^* \omega_{\mathcal{O}}^+, \tag{3.7}$$

where $\pi_2 \colon V^0 \times \mathcal{O} \to \mathcal{O}$ is projection on the second factor. This follows since for all $\xi \in \mathfrak{g}$,

$$\begin{aligned} \langle \mathbf{J}_{\mathcal{O}} \circ \chi(\alpha_q, \nu), \xi \rangle &= \langle \alpha_q + \mathcal{A}(q)^* \nu, \xi_Q(q) \rangle = \langle \alpha_q, \xi_Q(q) \rangle + \langle \mathcal{A}(q)^* \nu, \xi_Q(q) \rangle \\ &= 0 + \langle \nu, \xi \rangle, \end{aligned}$$

so that $\mathbf{J}_{\mathcal{O}} \circ \chi = \pi_2$.

THE FIRST TERM. The first term is a little more complicated and splits into a sum of terms. We begin by considering the pull back by χ of the restriction to $\mathbf{J}^{-1}(\mathcal{O})$ of the canonical one-form, and then we compute the exterior derivative of this one-form.

LEMMA 3.3. *We have*

$$\chi^* \iota_{\mathcal{O}}^* \Theta = \pi_1^* \iota_{V^0}^* \Theta + \varpi, \tag{3.8}$$

where $\iota_{V^0}: V^0 \to T^*Q$ is inclusion, Θ is the canonical one-form on T^*Q, $\pi_1: V^0 \times \mathcal{O} \to V^0$ is projection on the first factor, and $\varpi \in \Omega^1(V^0 \times \mathcal{O})$ is given by

$$\varpi(\alpha_q, \nu)(X_{\alpha_q}, X_\nu^{\xi'}) = \langle \nu, \mathcal{A}(q)(T_{\alpha_q}\tau_{T^*Q}(X_{\alpha_q}))\rangle \tag{3.9}$$

for $(X_{\alpha_q}, X_\nu^{\xi'}) \in T_{(\alpha_q,\nu)}(V^0 \times \mathcal{O})$, where $X_\nu^{\xi'} \in T_\nu\mathcal{O}$ denotes the infinitesimal generator for the left action of G on $\mathcal{O}$, $X_\nu^{\xi'} = -\operatorname{ad}_{\xi'}^* \nu$.

PROOF. Let $t \mapsto (\alpha_q(t), \operatorname{Ad}^*_{\exp -t\xi'} \nu)$ be a curve in $V^0 \times \mathcal{O}$, through the point (α_q, ν) such that $\frac{d}{dt}|_{t=0}\alpha_q(t) = X_{\alpha_q} \in T_{\alpha_q}V^0$. Since $\frac{d}{dt}|_{t=0} \operatorname{Ad}^*_{\exp -t\xi'} \nu = X_\nu^{\xi'}$, we get

$$\chi^* \iota_{\mathcal{O}}^* \Theta(\alpha_q, \nu)(X_{\alpha_q}, X_\nu^{\xi'}) = \iota_{\mathcal{O}}^*\Theta(\alpha_q + \mathcal{A}(q)^*\nu)(T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'})). \tag{3.10}$$

Now,

$$\begin{aligned} T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'}) &= \frac{d}{dt}\Big|_{t=0}\Big(\chi\big(\alpha_q(t), \operatorname{Ad}^*_{\exp -t\xi'} \nu\big)\Big) \\ &= \frac{d}{dt}\Big|_{t=0}\big(\alpha_q(t) + \mathcal{A}(q)^*(\operatorname{Ad}^*_{\exp -t\xi'} \nu)\big) \\ &= X_{\alpha_q} - \mathcal{A}(q)^*(\operatorname{ad}^*_{\xi'} \nu) \end{aligned}$$

where we use the fact that the curve $t \mapsto \mathcal{A}(q)^*(\operatorname{Ad}^*_{\exp -t\xi'} \nu)$ lies in the single fiber, T_q^*Q for all t. Thus, the right hand side in (3.10) becomes

$$\begin{aligned} &\iota_{\mathcal{O}}^*\Theta(\alpha_q + \mathcal{A}(q)^*\nu)(T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'}) \\ &= \Theta(\alpha_q + \mathcal{A}(q)^*\nu)\big(T_{(\alpha_q,\nu)}\chi(X_{\alpha_q}, X_\nu^{\xi'})\big) \\ &= \Theta(\alpha_q + \mathcal{A}(q)^*\nu)\big(X_{\alpha_q} - \mathcal{A}(q)^*(\operatorname{ad}^*_{\xi'} \nu)\big) \\ &= \langle \alpha_q + \mathcal{A}(q)^*\nu, T\tau_{T^*Q} \cdot X_{\alpha_q}\rangle \\ &= \langle \alpha_q, T\tau_{T^*Q} \cdot X_{\alpha_q}\rangle + \langle \nu, \mathcal{A}(q)(T\tau_{T^*Q} \cdot X_{\alpha_q})\rangle \\ &= \pi_1^*\iota_{V^0}^*\Theta + \varpi. \end{aligned}$$

The third equality holds because, for all t,

$$\tau_{T^*Q}\big(\alpha_q(t) + \mathcal{A}(q)^*(\operatorname{Ad}^*_{\exp -t\xi'} \nu)\big) = \tau_{T^*Q}\big(\alpha_q(t)\big). \qquad \blacksquare$$

Computing the exterior derivative,

$$\chi^* \iota_{\mathcal{O}}^*(-\mathbf{d}\Theta) = -\mathbf{d}(\pi_1^* \iota_{V^0}^* \Theta + \varpi) = \pi_1^* \iota_{V^0}^* \Omega - \mathbf{d}\varpi,$$

so that

$$\chi^*(\pi_{\mathcal{O}}^* \Omega_{\mathcal{O}}) = \chi^* \iota_{\mathcal{O}}^* \Omega - \chi^* \mathbf{J}_{\mathcal{O}}^* \omega_{\mathcal{O}}^+ = \pi_1^* \iota_{V^0}^* \Omega - \mathbf{d}\varpi - \pi_2^* \omega_{\mathcal{O}}^+.$$

The form ϖ is $\varpi = (\tau_{T^*Q} \times \mathrm{id})^* \alpha$, the pull back to $V^0 \times \mathcal{O}$ of the one-form α on $Q \times \mathcal{O}$ defined by

$$\alpha(q, \nu)(X_q, X_\nu) = \langle \nu, \mathcal{A}(q)(X_q) \rangle. \tag{3.11}$$

We are implicitly restricting the domain of τ_{T^*Q} to the sub-bundle V^0.

3.2. *Computation of* $\mathbf{d}\alpha$. The philosophy of the computation will be to make use of the connection to decompose tangent vectors to Q in terms of their horizontal and vertical parts. Of course we expect the curvature of the connection to appear in the resulting formula. However the presence of the pairing with ν, which varies over the coadjoint orbit $\mathcal{O}$ must be dealt with carefully.

We begin with an elementary but useful fact concerning the Jacobi-Lie bracket of vector fields on the cartesian product of two manifolds.

LEMMA 3.4. *Let M and N be two smooth manifolds of dimension m and n respectively and consider their Cartesian product $M \times N$. Suppose we have two vector fields (X^M, X^N) and (Y^M, Y^N) on $M \times N$, each with the property that the tangent vector to M is independent of N, and that the tangent vector to N is independent of M. Then, the Jacobi Lie bracket of these two vector fields is also of this type. In fact, we have,*

$$[(X^M, X^N), (Y^M, Y^N)] = ([X^M, Y^M]_M, [X^N, Y^N]_N). \tag{3.12}$$

This is readily proved using the local coordinate expression of the bracket.

To determine $\mathbf{d}\alpha \in \Omega^2(Q \times \mathcal{O})$, it suffices, by bilinearity and skew symmetry, to compute its value on pairs of tangent vectors to Q of the type

- hor_q, hor_q
- hor_q, ver_q
- ver_q, ver_q.

To carry this out, we will extend each tangent vector to be horizontal or vertical in an entire neighborhood of the point in question and use the fact that $\mathbf{d}\alpha$ is a tensor.

CASE 1. $X_q, Y_q \in \mathrm{Hor}_q Q$. We consider $(X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}) \in T_{(q,\nu)}(Q \times \mathcal{O})$. Extend X_q to the horizontal vector field $\tilde{X}_Q \in \mathrm{Hor}\, Q$ and similarly extend Y_q to $\tilde{Y}_Q$. We extend the second components of each tangent vector in the obvious way to be infinitesimal generators of the given Lie algebra element. That is we extend $X_\nu^{\xi'}$ to $\xi'_{\mathcal{O}}$ and similarly for $Y_\nu^{\eta'}$. Denote by $\tilde{X}$, the extended vector field on a neighborhood of $Q \times \mathcal{O}$ given by $(\tilde{X}_Q, \xi'_{\mathcal{O}})$, and similarly for $\tilde{Y}$. We then have

$$\begin{aligned} &\mathbf{d}\alpha(q,\nu)\big((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})\big) \\ (3.13) \qquad &= (X_q, X_\nu^{\xi'}) \cdot \alpha(\tilde{Y}_Q, \eta'_{\mathcal{O}}) - (Y_q, Y_\nu^{\eta'}) \cdot \alpha(\tilde{X}_Q, \xi'_{\mathcal{O}}) - \alpha([\tilde{X}, \tilde{Y}])(q,\nu). \end{aligned}$$

Notice that the first term vanishes since, if we take a curve $t \mapsto \big(q(t), \nu(t)\big)$ through the point (q,ν) such that $\big(\dot{q}(0), \dot{\nu}(0)\big) = (X_q, X_\nu^{\xi'})$, we have

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} \alpha\big(q(t),\nu(t)\big)\Big(\tilde{Y}_Q\big(q(t)\big), \eta'_{\mathcal{O}}\big(\nu(t)\big)\Big) &= \frac{d}{dt}\Big|_{t=0} \langle \nu(t), \mathcal{A}(q(t)) \cdot \tilde{Y}_Q\big(q(t)\big)\rangle \\ (3.14) \qquad &= 0 \end{aligned}$$

since for all t, $\tilde{Y}_Q\big(q(t)\big) \in \operatorname{Hor}_{q(t)} Q$. Similarly, the second term vanishes. Now, by Lemma 3.4, we have $[\tilde{X}, \tilde{Y}]_{Q \times \mathcal{O}} = ([\tilde{X}_Q, \tilde{Y}_Q]_Q, [\xi'_{\mathcal{O}}, \eta'_{\mathcal{O}}]_{\mathcal{O}})$ leaving

$$(3.15) \qquad \mathbf{d}\alpha(q,\nu)\big((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})\big) = -\big\langle \nu, \mathcal{A}(q)\big([\tilde{X}_Q, \tilde{Y}_Q](q)\big)\big\rangle.$$

However, since $\tilde{X}_Q$ and $\tilde{Y}_Q$ are horizontal vector fields, it follows that

$$(3.16) \qquad \mathcal{A}([\tilde{X}_Q, \tilde{Y}_Q]) = -\operatorname{Curv}_{\mathcal{A}}(\tilde{X}_Q, \tilde{Y}_Q)$$

so that

$$(3.17) \qquad \mathbf{d}\alpha(q,\nu)\big((X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'})\big) = \langle \nu, \operatorname{Curv}_{\mathcal{A}}(X_q, Y_q)\rangle.$$

CASE 2. $X_q \in \operatorname{Hor}_q Q$, $Y_q \in \operatorname{Ver}_q Q$. Using the same notation for vector fields as in the previous case, we let $\tilde{X}_Q$ denote the horizontal vector field extending X_q. Let $\eta = \mathcal{A}(q)(Y_q)$. Since Y_q is vertical we have $\eta_Q(q) = Y_q$. Then η_Q is a vertical extension of Y_q. With these extensions, we have

$$\begin{aligned} &\mathbf{d}\alpha(q,\nu)(X_q, X_\nu^{\xi'}), (Y_q, Y_\nu^{\eta'}) \\ (3.18) \qquad &= (X_q, X_\nu^{\xi'}) \cdot \alpha(\eta_Q, \eta'_{\mathcal{O}}) - (Y_q, Y_\nu^{\eta'}) \cdot \alpha\big((\tilde{X}_Q, \xi'_{\mathcal{O}})\big) - \alpha(q,\nu)([\tilde{X}, \tilde{Y}]). \end{aligned}$$

Consider the first term. Let $t \mapsto \big(q(t), \nu(t)\big)$ be a curve through (q,ν) with $\big(\dot{q}(0), \dot{\nu}(0)\big) = (X_q, -\operatorname{ad}^*_{\xi'} \nu)$. Then

$$\begin{aligned} (X_q, X_\nu^{\xi'}) \cdot \alpha(\eta_Q, \eta'_{\mathcal{O}}) &= \frac{d}{dt}\Big|_{t=0} \alpha\big(q(t), \nu(t)\big)\Big(\eta_Q\big(q(t)\big), \eta'_{\mathcal{O}}\big(\nu(t)\big)\Big) \\ &= \frac{d}{dt}\Big|_{t=0} \Big\langle \nu(t), \mathcal{A}\big(q(t)\big)\Big(\eta_Q\big(q(t)\big)\Big)\Big\rangle \\ &= \frac{d}{dt}\Big|_{t=0} \langle \nu(t), \eta\rangle = \langle -\operatorname{ad}^*_{\xi'} \nu, \eta\rangle. \end{aligned}$$

The second term vanishes since $\alpha(\tilde{X}_Q, \xi'_{\mathcal{O}}) = 0$ for $\tilde{X}_Q \in \operatorname{Hor} Q$. Recall that for $\tilde{X}_Q$ a horizontal vector field, we have, for all $\eta \in \mathfrak{g}$,

$$(3.19) \qquad [\tilde{X}_Q, \eta_Q] \in \operatorname{Hor} Q.$$

This fact, together with Lemma 3.4, gives

$$\alpha(q,\nu)([\tilde{X},\tilde{Y}]) = \langle \nu, \mathcal{A}(q)([\tilde{X}_Q, \eta_Q])\rangle = 0, \tag{3.20}$$

so that

$$\mathbf{d}\alpha(q,\nu)(X_q, X_\nu^{\xi'}),(Y_q, Y_\nu^{\eta'}) = \langle -\operatorname{ad}^*_{\xi'}\nu, \eta\rangle. \tag{3.21}$$

CASE 3. $X_q, Y_q \in \operatorname{Ver}_q Q$. Let $\xi = \mathcal{A}(q)(X_q)$ and $\eta = \mathcal{A}(q)(Y_q)$. We choose extensions to be vertical globally. Thus, $\tilde{X}_Q = \xi_Q$ and $\tilde{Y}_Q = \eta_Q$. Then we compute each term in the expression for $\mathbf{d}\alpha$.

The first term will again be

$$\langle -\operatorname{ad}^*_{\xi'}\nu, \eta\rangle \tag{3.22}$$

since $\alpha(q,\nu)(\eta_Q, \eta'_{\mathcal{O}}) = \langle \nu, \eta\rangle$, *i.e.*, $\iota_{(\eta_Q,\eta'_{\mathcal{O}})}\alpha \colon Q\times\mathcal{O}\to\mathbb{R}$ is independent of Q.

The second term is computed similarly to be $\langle \operatorname{ad}^*_{\eta'}\nu, \xi\rangle$. For the last term, recall that for left actions, $G\times Q\to Q$, we have

$$[\xi_Q, \eta_Q] = -[\xi,\eta]_Q \tag{3.23}$$

so that

$$\alpha(q,\nu)([\tilde{X},\tilde{Y}]) = \langle \nu, \mathcal{A}(q)([\xi_Q,\eta_Q])\rangle = -\langle \nu, \mathcal{A}(q)([\xi,\eta]_Q)(q)\rangle = -\langle \nu, [\xi,\eta]\rangle.$$

Therefore,

$$\mathbf{d}\alpha(q,\nu)(X_q, X_\nu^{\xi'}),(Y_q, Y_\nu^{\eta'}) = \langle \nu, [\eta,\xi']\rangle + \langle \nu, [\eta',\xi]\rangle + \langle \nu, [\xi,\eta]\rangle. \tag{3.24}$$

We now collect these results to obtain a formula for the two form relative to a decomposition of the tangent vectors to Q into their horizontal and vertical projections.

THEOREM 3.5. *Let* $(X_q, X_\nu^{\xi'}),(Y_q, Y_\nu^{\eta'}) \in T_{(q,\nu)}(Q\times\mathcal{O})$. *Let*

$$\xi = \mathcal{A}(q)(X_q) \quad \textit{and} \quad \eta = \mathcal{A}(q)(Y_q)$$

so that

$$X_q = \xi_Q(q) + \operatorname{Hor}_q X_q, \quad Y_q = \eta_Q(q) + \operatorname{Hor}_q Y_q \tag{3.25}$$

where Hor_q *denotes the horizontal projection onto the horizontal distribution. We then have*

$$\begin{aligned}&\mathbf{d}\alpha(q,\nu)\big((X_q, X_\nu^{\xi'}),(Y_q, Y_\nu^{\eta'})\big)\\ &\quad= \langle \nu, [\eta',\xi]\rangle + \langle \nu, [\eta,\xi']\rangle + \langle \nu, [\xi,\eta]\rangle + \langle \nu, \operatorname{Curv}_{\mathcal{A}}(q)(X_q, Y_q)\rangle.\end{aligned} \tag{3.26}$$

PROOF. The proof is a straightforward computation:

$$
\begin{aligned}
&\mathbf{d}\alpha(q,\nu)\big((X_q, X_\nu^{\xi'}),(Y_q, Y_\nu^{\eta'})\big)\\
&= \mathbf{d}\alpha(q,\nu)\Big(\big(\xi_Q(q)+\operatorname{Hor}_q X_q, X_\nu^{\xi'}\big),\big(\eta_Q(q)+\operatorname{Hor}_q Y_q, Y_\nu^{\eta'}\big)\Big)\\
&= \mathbf{d}\alpha(q,\nu)\Big(\Big(\xi_Q(q),\frac{1}{2}X_\nu^{\xi'}\Big)+\Big(\operatorname{Hor}_q X_q,\frac{1}{2}X_\nu^{\xi'}\Big),\Big(\eta_Q(q),\frac{1}{2}Y_\nu^{\eta'}\Big)\\
&\qquad +\Big(\operatorname{Hor}_q Y_q,\frac{1}{2}Y_\nu^{\eta'}\Big)\Big)\\
&= \mathbf{d}\alpha(q,\nu)\Big(\Big(\xi_Q(q),\frac{1}{2}X_\nu^{\xi'}\Big),\Big(\eta_Q(q),\frac{1}{2}Y_\nu^{\eta'}\Big)\Big)\\
&\quad + \mathbf{d}\alpha(q,\nu)\Big(\Big(\xi_Q(q),\frac{1}{2}X_\nu^{\xi'}\Big),\Big(\operatorname{Hor}_q Y_q,\frac{1}{2}Y_\nu^{\eta'}\Big)\Big)\\
&\quad + \mathbf{d}\alpha(q,\nu)\Big(\Big(\operatorname{Hor}_q X_q,\frac{1}{2}X_\nu^{\xi'}\Big),\Big(\eta_Q(q),\frac{1}{2}Y_\nu^{\eta'}\Big)\Big)\\
&\quad + \mathbf{d}\alpha(q,\nu)\Big(\Big(\operatorname{Hor}_q X_q,\frac{1}{2}X_\nu^{\xi'}\Big),\Big(\operatorname{Hor}_q Y_q,\frac{1}{2}Y_\nu^{\eta'}\Big)\Big)\\
&= \Big\langle \nu,\Big[\frac{1}{2}\eta',\xi\Big]\Big\rangle + \langle \nu,[\xi,\eta]\rangle + \langle \nu,[\xi,\eta]\rangle + \Big\langle \nu,\Big[\eta,\frac{1}{2}\xi'\Big]\Big\rangle\\
&\quad - \langle \operatorname{ad}^*_{\frac{1}{2}\eta'}\nu,\xi\rangle + \langle -\operatorname{ad}^*_{\frac{1}{2}\xi'}\nu,\eta\rangle + \langle \nu,\operatorname{Curv}_{\mathcal{A}}(X_q,Y_q)\rangle\\
&= \langle \nu,[\eta',\xi]\rangle + \langle \nu,[\eta,\xi']\rangle + \langle \nu,[\xi,\eta]\rangle\\
&\quad + \langle \nu,\operatorname{Curv}_{\mathcal{A}}(q)(X_q,Y_q)\rangle
\end{aligned}
$$

where we have used the relations determined in the previous section. ∎

4. The reduced form. Recall that

$$
\begin{aligned}
\chi^*(\pi_{\mathcal{O}}^*\Omega_{\mathcal{O}}) &= \chi^*\iota_{\mathcal{O}}^*\Omega - \chi^*\mathbf{J}_{\mathcal{O}}^*\omega_{\mathcal{O}}^+\\
&= \pi_1^*\iota_{V^0}^*\Omega - \mathbf{d}\varpi - \pi_2^*\omega_{\mathcal{O}}^+\\
&= \pi_1^*\iota_{V^0}^*\Omega - (\tau_{T^*Q}\times \mathrm{id})^*\mathbf{d}\alpha - \pi_2^*\omega_{\mathcal{O}}^+. \qquad (4.1)
\end{aligned}
$$

We have already established the G-invariance of this form. Notice that the first term is independently G-invariant since, if we denote the action of G on $V^0\times\mathcal{O}$ by $\psi^{V^0\times\mathcal{O}}$ and the action of G on T^*Q by ψ, we have

$$
\begin{aligned}
(\psi_g^{V^0\times\mathcal{O}})^*\pi_1^*\iota_{V^0}^*\Omega &= \pi_1^*\psi_g^*\iota_{V^0}^*\Omega = \pi_1^*\iota_{V^0}^*\psi_g^*\Omega\\
&= \pi_1^*\iota_{V^0}^*\Omega
\end{aligned}
$$

since $\pi_1\circ\psi_g^{V^0\times\mathcal{O}}(\alpha_q,\nu) = g\cdot\alpha_q = \psi_g\circ\pi_1(\alpha_q,\nu)$. Thus, the sum of the last two terms is G invariant. Furthermore, the G invariance of the last two terms as forms on $V^0\times\mathcal{O}$, is really G invariance of a form on $Q\times\mathcal{O}$ since

$$
\begin{aligned}
(\tau_{T^*Q}\times\mathrm{id})\circ\psi_g^{V^0\times\mathcal{O}}(\alpha_q,\nu) &= (\tau_{T^*Q}\times\mathrm{id})(g\cdot\alpha_q, g\cdot\nu)\\
&= (g\cdot q, g\cdot\nu) = \psi_g^{Q\times\mathcal{O}}(q,\nu)\\
&= \psi_g^{Q\times\mathcal{O}}\circ(\tau_{T^*Q}\times\mathrm{id})(\alpha_q,\nu).
\end{aligned}
$$

4.1. *The part that drops to $\tilde{\mathcal{O}}$.* We begin with the proof of the vanishing of the two-form $\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+$ on vertical vectors.

PROPOSITION 4.1. *The two form, $\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+$ on $Q \times \mathcal{O}$ vanishes on vertical vectors of the bundle $Q \times \mathcal{O} \to \tilde{\mathcal{O}}$. It therefore uniquely determines a two form on $\tilde{\mathcal{O}}$.*

PROOF. For the two form $\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+$ on $Q \times \mathcal{O}$ to drop to the quotient, $\tilde{\mathcal{O}}$, we must have both G invariance of the form and also the property that it vanish on the vertical fibers. To see this, fix $\xi \in \mathfrak{g}$ and let $(Y_q, Y_\nu^{\eta'}) \in T_{(q,\nu)}(Q \times \mathcal{O})$. As usual, let $\eta = \mathcal{A}(q)(Y_q)$. Since the action of G on $Q \times \mathcal{O}$ is the diagonal action, we have

$$\xi_{Q\times\mathcal{O}}(q,\nu) = \frac{d}{dt}\bigg|_{t=0}(\exp t\xi \cdot q, \mathrm{Ad}^*_{\exp -t\xi}\,\nu) = \big(\xi_Q(q), X^\xi_\nu\big). \tag{4.2}$$

We then have

$$\begin{aligned}
&(\mathbf{d}\alpha + \pi_2^*\omega_{\mathcal{O}}^+)(q,\nu)\Big(\big(\xi_Q(q), X^\xi_\nu\big), (Y_q, Y_\nu^{\eta'})\Big)\\
&\quad = \mathbf{d}\alpha(q,\nu)\Big(\big(\xi_Q(q), X^\xi_\nu\big), (Y_q, Y_\nu^{\eta'})\Big) + \omega_{\mathcal{O}}^+(-\operatorname{ad}^*_\xi \nu, -\operatorname{ad}^*_{\eta'}\nu)\\
&\quad = \langle \nu, [\eta',\xi]\rangle + \langle \nu, [\eta,\xi]\rangle + \langle \nu, [\xi,\eta]\rangle + \omega_{\mathcal{O}}^+(\operatorname{ad}^*_\xi \nu, \operatorname{ad}^*_{\eta'}\nu)\\
&\quad = \langle \nu, [\eta',\xi]\rangle + \langle \nu, [\xi,\eta']\rangle = 0.
\end{aligned}$$

Notice that the curvature term in the formula for $\mathbf{d}\alpha$ vanishes since it is evaluated on a vertical vector $\xi_Q(q)$. ∎

4.2. *The part that drops to $T^*(Q/G)$.* We now characterize the first term of

$$\pi_1^*\iota_{V^0}^*\Omega - (\tau_{T^*Q} \times \mathrm{id})^*\mathbf{d}\alpha - \pi_2^*\omega_{\mathcal{O}}^+ \tag{4.3}$$

as the pull back relative to Γ of the canonical form on $T^*(Q/G)$.

PROPOSITION 4.2. *Denote by*

$$\pi_{\mathcal{A}}\colon V^0 \to T^*(Q/G) \tag{4.4}$$

the map given by $\alpha_q \mapsto \mathrm{hor}_q^*\,\alpha_q \in T^*_{[q]}(Q/G)$. Note that this is simply the map Γ restricted to V^0. Let Θ denote the canonical one-form on T^*Q and $\Theta_{Q/G}$ the canonical one-form on $T^*(Q/G)$. We then have

$$\pi_{\mathcal{A}}^*\Theta_{Q/G} = \iota_{V^0}^*\Theta \tag{4.5}$$

from which it follows that

$$\pi_{\mathcal{A}}^*\Omega_{Q/G} = \iota_{V^0}^*\Omega. \tag{4.6}$$

PROOF. Let $X_{\alpha_q} \in T_{\alpha_q} V^0$. We have

$$\begin{aligned}\pi_{\mathcal{A}}^* \Theta_{Q/G}(X_{\alpha_q}) &= \Theta_{Q/G}(\mathrm{hor}_q^* \alpha_q) \\ &= \langle \mathrm{hor}_q^* \alpha_q, T\tau_{Q/G} \circ T\pi_{\mathcal{A}} \cdot X_{\alpha_q} \rangle.\end{aligned}$$

We need to compute the derivative of the composition,

$$\tau_{Q/G} \circ \pi_{\mathcal{A}} : V^0 \to Q/G. \tag{4.7}$$

Let $t \mapsto \alpha_q(t) \in V^0$ be a smooth curve through α_q such that $\dot{\alpha}_q(0) = X_{\alpha_q}$. Let $q(t) = \tau_{T^* Q}\big(\alpha_q(t)\big)$. Then

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} \tau_{Q/G} \circ \pi_{\mathcal{A}}\big(\alpha_q(t)\big) &= \left.\frac{d}{dt}\right|_{t=0} \tau_{Q/G}\Big(\mathrm{hor}_{q(t)}^* \big(\alpha_q(t)\big)\Big) \\ &= \left.\frac{d}{dt}\right|_{t=0} [q(t)] = T\pi \circ T\tau_{T^* Q} \cdot X_{\alpha_q}.\end{aligned}$$

Thus,

$$\langle \mathrm{hor}_q^* \alpha_q, T\tau_{Q/G} \circ T\pi_{\mathcal{A}} \cdot X_{\alpha_q} \rangle = \langle \alpha_q, \mathrm{hor}_q \circ T\pi \circ T\tau_{T^* Q} \cdot X_{\alpha_q} \rangle. \tag{4.8}$$

On the other hand, we have

$$\begin{aligned}\iota_{V^0}^* \Theta(\alpha_q)(X_{\alpha_q}) &= \langle \alpha_q, T\tau_{T^* Q} \cdot X_{\alpha_q} \rangle \\ &= \langle \alpha_q, \mathrm{Hor}_q T\tau_{T^* Q} \cdot X_{\alpha_q} + \mathrm{Ver}_q T\tau_{T^* Q} \cdot X_{\alpha_q} \rangle \\ &= \langle \alpha_q, \mathrm{Hor}_q T\tau_{T^* Q} \cdot X_{\alpha_q} \rangle \\ &= \langle \alpha_q, \mathrm{hor}_q \circ T\pi \circ T\tau_{T^* Q} \cdot X_{\alpha_q} \rangle\end{aligned}$$

where the third equality follows from the fact that α_q annihilates vertical vectors. ■

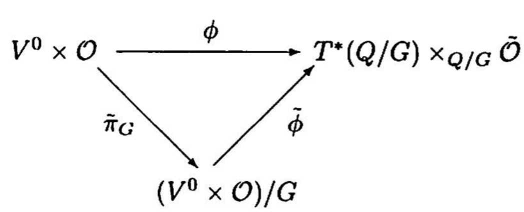
4.3. *A final piece of diagram chasing.* Recall that we have the following maps:

$$V^0 \times \mathcal{O} \xrightarrow{\tau_{T^* Q} \times \mathrm{id}} Q \times \mathcal{O} \xrightarrow{\pi_G} \tilde{\mathcal{O}}.$$

Define the map ϕ as follows:

$$\phi(\alpha_q, \nu) = \big(\mathrm{hor}_q^* \circ \pi_1, \pi_G \circ (\tau_{T^* Q} \times \mathrm{id})\big). \tag{4.9}$$

It is easy to see that ϕ is G-invariant, so that we have the following commutative diagram.



It is straightforward to check that the map $\tilde{\phi}$ is invertible and therefore determines a bundle isomorphism.

THEOREM 4.3. *Denote by $\omega_{\rm red}$ the reduced symplectic form on the symplectic reduced space $T^*(Q/G) \times_{Q/G} \tilde{\mathcal{O}}$. We then have the formula*

$$\omega_{\rm red} = \Omega_{Q/G} - \beta$$

where β is the unique two form on $\tilde{\mathcal{O}}$ determined by

$$\pi_G^* \beta = \mathbf{d}\alpha + \pi_2^* \omega_{\mathcal{O}}^+$$

and, as in Theorem 3.5,

$$\begin{aligned} &\mathbf{d}\alpha(q,\nu)\big((X_q, X_\nu^{\xi'}),(Y_q, Y_\nu^{\eta'})\big) \\ &= \langle \nu, [\eta', \xi]\rangle + \langle \nu, [\eta, \xi']\rangle + \langle \nu, [\xi, \eta]\rangle + \langle \nu, \mathrm{Curv}_{\mathcal{A}}(q)(X_q, Y_q)\rangle. \end{aligned} \tag{4.10}$$

PROOF. The two-form $\omega_{\rm red}$ is the unique two-form on $T^*(Q/G) \times_{Q/G} \tilde{\mathcal{O}}$ such that $\tilde{\phi}^* \omega_{\rm red} = \bar{\chi}^* \Omega_{\mathcal{O}}$, where $\bar{\chi}^* \Omega_{\mathcal{O}}$ (see equation (3.5)) is the unique two-form on $(V^0 \times \mathcal{O})/G$ such that

$$\tilde{\pi}_G^* \bar{\chi}^* \Omega_{\mathcal{O}} = \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+.$$

We then have

$$\tilde{\pi}_G^* \tilde{\phi}^* \omega_{\rm red} = \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+.$$

However, since $\tilde{\phi} \circ \tilde{\pi}_G = \phi$, we have

$$\phi^* \omega_{\rm red} = \pi_1^* \iota_{V^0}^* \Omega - (\tau_{T^*Q} \times \mathrm{id})^* \mathbf{d}\alpha - \pi_2^* \omega_{\mathcal{O}}^+ \tag{4.11}$$

from which we can read off $\omega_{\rm red}$:

$$\begin{aligned} &\phi^* \omega_{\rm red}(\alpha_q, \nu)\big((X_{\alpha_q}, X_\nu),(Y_{\alpha_q}, Y_\nu)\big) \\ &\quad = \omega_{\rm red}(\mathrm{hor}_q^* \alpha_q, [q,\nu]) \\ &\qquad \times \Big(\big(T(\mathrm{hor}^* \circ \pi_1)(X_{\alpha_q}, X_\nu), T(\pi_G \circ \tau_{T^*Q} \times \mathrm{id})(X_{\alpha_q}, X_\nu)\big), \\ &\qquad\quad \big(T(\mathrm{hor}^* \circ \pi_1)(Y_{\alpha_q}, Y_\nu), T(\pi_G \circ \tau_{T^*Q} \times \mathrm{id})(Y_{\alpha_q}, Y_\nu)\big)\Big). \end{aligned}$$

Note that $T(\mathrm{hor}^* \circ \pi_1)(X_{\alpha_q}, X_\nu) = T\pi_{\mathcal{A}} X_{\alpha_q}$ and

$$T\big(\pi_G \circ (\tau_{T^*Q} \times \mathrm{id})\big)(X_{\alpha_q}, X_\nu) = T_{(q,\nu)}\pi_G \cdot (T\tau_{T^*Q} X_{\alpha_q}, X_\nu).$$

The right hand side of equation (4.11) becomes

$$\begin{aligned} &\Omega(\alpha_q)(X_{\alpha_q}, Y_{\alpha_q}) - (\tau_{T^*Q} \times \mathrm{id})^* \pi_G^* \beta(X_{\alpha_q}, Y_{\alpha_q}) \\ &\quad = \Omega_{Q/G}\big(\pi_{\mathcal{A}}(\alpha_q)\big)(T\pi_{\mathcal{A}} X_{\alpha_q}, T\pi_{\mathcal{A}} Y_{\alpha_q}) \\ &\qquad - \beta([q,\nu])\big(T_{(q,\nu)}\pi_G \cdot (T\tau_{T^*Q} X_{\alpha_q}, X_\nu), T_{(q,\nu)}\pi_G \cdot (T\tau_{T^*Q} Y_{\alpha_q}, Y_\nu)\big), \end{aligned}$$

from which the the claim follows. ∎

5. The extreme cases. The obvious extreme cases are $Q = G$ and G Abelian. We first consider the case $Q = G$. Then, Q/G reduces to a point, and the associated bundle is simply the coadjoint orbit through ν_0. $\tilde{\mathcal{O}} = Q \times \mathcal{O}/G = G \times \mathcal{O}/G \simeq \mathcal{O}$. Consider a tangent vector to the coadjoint orbit through a point ν, given by $-\operatorname{ad}^*_{\xi'}\nu$. Represent this tangent vector with the curve through ν, $t \mapsto \operatorname{Ad}^*_{\exp -t\xi'}\nu$. We must find a lift to $G \times \mathcal{O}$ of such a tangent vector. The projection, $\pi_G: G \times \mathcal{O} \to \mathcal{O}$ is given by $(g, \nu) \mapsto g^{-1}\nu$ since $[g, \nu] = [e, g^{-1}\nu]$. More generally, consider a curve through $(e, \nu) \in G \times \mathcal{O}$ denoted by $t \mapsto (g(t), \nu(t))$. Let $\xi = \dot{g}(0)$. Since $\mathcal{A}(e)(\dot{g}(0)) = \dot{g}(0)$, this is consistent notation. The projection of this curve to $\mathcal{O}$ is given by

$$\pi_G(g(t), \nu(t)) = g(t)^{-1}\nu(t) = \operatorname{Ad}^*_{g(t)}\nu(t) \tag{5.1}$$

and therefore we require

$$\frac{d}{dt}\bigg|_{t=0} \operatorname{Ad}^*_{g(t)}\nu(t) = -\operatorname{ad}^*_{\xi'}\nu,$$

which implies

$$\operatorname{Ad}^*_g \dot{\nu}(o) + \operatorname{ad}^*_{\dot{g}(0)}\nu = -\operatorname{ad}^*_{\xi'}\nu, \tag{5.2}$$

from which it follows that

$$\dot{\nu}(0) = -\operatorname{ad}^*_{\xi'}\nu - \operatorname{ad}^*_{\xi}\nu.$$

Equations (4.1) and (4.10) give

$$\begin{aligned}
&\omega_{\text{red}}(\nu)(-\operatorname{ad}^*_{\xi'}\nu, \operatorname{ad}^*_{\eta'}\nu)\\
&= \Omega_{Q/G} - \beta(e,\nu)\big((\xi, X^{\xi'+\xi}_\nu), (\eta, Y^{\eta'+\eta}_\nu)\big)\\
&= -(\mathbf{d}\alpha + \pi_2^*\omega^+_{\mathcal{O}})(e,\nu)\big((\xi, X^{\xi'+\xi}_\nu), (\eta, Y^{\eta'+\eta}_\nu)\big)\\
&= -\big(\langle \nu, [\eta' + \eta]\rangle + \langle \nu, [\eta, \xi' + \xi]\rangle + \langle \nu, [\xi, \eta]\rangle\\
&\qquad + \operatorname{Curv}_{\mathcal{A}}(e)(\xi, \eta) + \langle \nu, [\xi' + \xi, \eta' + \eta]\rangle\big)\\
&= -\langle \nu, [\xi', \eta']\rangle,
\end{aligned}$$

where the last equality follows from the fact that the curvature term vanishes on vertical vectors and an expansion of the Lie algebra brackets.

For G Abelian, the fibers of the $\tilde{\mathcal{O}}$ bundle collapse and we are left with just $T^*(Q/G)$. The reduced symplectic form, from equations (4.1) and (4.10) is then

$$\omega_{\text{red}} = \Omega - \langle \nu, \operatorname{Curv}_{\mathcal{A}}\rangle \tag{5.3}$$

since all brackets vanish.

ACKNOWLEDGEMENTS. We thank Anthony Blaom, Hernan Cendra, Sameer Jalnapurkar, Gerard Misiolek and Tudor Ratiu for helpful comments and inspiration.

REFERENCES

1. R. Abraham and J. E. Marsden, *Foundations of Mechanics*. Second Edition, Addison-Wesley, 1978.
2. J. M. Arms, J. E. Marsden and V. Moncrief, *Symmetry and bifurcations of momentum mappings*. Comm. Math. Phys. **78**(1981), 455–478.
3. ———, *The structure of the space solutions of Einstein's equations: II Several Killings fields and the Einstein-Yang-Mills equations*. Ann. of Phys. **144**(1982), 81–106.
4. V. I. Arnold, *Sur la géométrie differentielle des groupes de Lie de dimenson infinie et ses applications à l'hydrodynamique des fluids parfaits*. Ann. Inst. Fourier (Grenoble) **16**(1966), 319–361.
5. ———, *Dynamical Systems III*. Encyclopedia of Mathematics **3**, Springer-Verlag, 1988.
6. ———, *Mathematical Methods of Classical Mechanics*. 2nd edition, Graduate Texts in Math. **60**, Springer-Verlag, 1989.
7. A. Blaom, *On geometric and dynamic phases*. Preprint, 1998.
8. ———, *Solution reconstruction in mechanical systems with symmetry*. Preprint, 1999; see http://www.alumni.caltech.edu/~ant/.
9. A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden and T. S. Ratiu, *The Euler-Poincaré equations and double bracket dissipation*. Comm. Math. Phys. **175**(1996), 1–42.
10. A. I. Bobenko, A. G. Reyman, and M. A. Semenov-Tian-Shansky, *The Kowalewski Top 99 years later: A Lax pair, generalizations and explicit solutions*. Comm. Math. Phys. **122**(1989), 321–354.
11. H. Cendra, H., D. D. Holm, M. J. W. Hoyle and J. E. Marsden, *The Maxwell-Vlasov equations in Euler-Poincaré form*. J. Math. Phys. **39**(1998), 3138–3157
12. H. Cendra, D. D. Holm, J. E. Marsden and T. S. Ratiu, *Lagrangian Reduction, the Euler–Poincaré Equations, and Semidirect Products*. Amer. Math. Soc. Transl. **186**(1998), 1–25.
13. H. Cendra, A. Ibort and J. E. Marsden, *Variational principal fiber bundles: a geometric theory of Clebsch potentials and Lin constraints*. J. Geom. Phys. **4**(1987), 183–206.
14. H. Cendra and J. E. Marsden, *Lin constraints, Clebsch potentials and variational principles*. Phys. D. **27**(1987), 63–89.
15. H. Cendra, J. E. Marsden and T. S. Ratiu, *Lagrangian reduction by stages*. Mem. Amer. Math. Soc., 2000, to appear.
16. R. Cushman and J. Śniatycki, *Hamiltonian mechanics on principal bundles*. C. R. Math. Rep. Acad. Sci. Canada, 2000, to appear.
17. A. Guichardet, *On rotation and vibration motions of molecules*. Ann. Inst. H. Poincaré **40**(1984), 329–342.
18. V. Guillemin and S. Sternberg, *The moment map and collective motion*. Ann. of Phys. **1278**(1980), 220–253.
19. ———, *Symplectic Techniques in Physics*. Cambridge University Press, 1984.
20. D. D. Holm, J. E. Marsden and T. S. Ratiu, *The Euler-Poincaré equations and semidirect products with applications to continuum theories*. Adv. in Math. **137**(1998), 1–81.
21. ———, *The Euler-Poincaré equations in geophysical fluid dynamics*. Proc. of the Isaac Newton Institute Programme on the Mathematics of Atmospheric and Ocean Dynamics, Cambridge University Press, 1999, to appear.
22. P. J. Holmes and J. E. Marsden, *Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups*. Indiana Univ. Math. J. **32**(1983), 273–310.
23. T. Iwai, *A geometric setting for classical molecular dynamics*. Ann. Inst. Henri Poincaré Phys. Théor. **47**(1987), 199–219.
24. S. M. Jalnapurkar and J. E. Marsden, *Reduction of Hamilton's Variational Principle*. Dynamics Stability Systems, 2000, to appear.
25. D. Kazhdan, B. Kostant and S. Sternberg, *Hamiltonian group actions and dynamical systems of Calogero type*. Comm. Pure Appl. Math. **31**(1978), 481–508.
26. A. A. Kirillov, *Unitary representations of nilpotent Lie groups*. Russian Math. Surveys **17**(1962), 53–104.
27. ———, *Elements of the Theory of Representations*. Grundlehren Math. Wiss. **220**, Springer-Verlag, 1976.

28. W. S. Koon and J. E. Marsden, *Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction.* SIAM J. Control Optim. **35**(1997), 901–929.
29. M. Kummer, *On the construction of the reduced phase space of a Hamiltonian system with symmetry.* Indiana Univ. Math. J. **30**(1981), 281–291.
30. N. E. Leonard and J. E. Marsden, *Stability and drift of underwater vehicle dynamics: mechanical systems with rigid motion symmetry.* Phys. D **105**(1997), 130–162.
31. S. Lie, *Theorie der Transformationsgruppen, Zweiter Abschnitt.* Teubner, Leipzig, 1890.
32. J. E. Marsden, *Lectures on Geometric Methods in Mathematical Physics.* SIAM, Philadelphia, PA, 1981.
33. ———, *Lectures on Mechanics.* London Math. Soc. Lecture Note Ser. **174**, Cambridge University Press, 1992.
34. ———, *Park City Lectures on mechanics, dynamics and symmetry.* Symplectic Geometry and Topology, IAS/Park City Math. Ser. **7**(1999), 335–430.
35. J. E. Marsden, G. Misiolek, M. Perlmutter and T. S. Ratiu, *Symplectic reduction for semidirect products and central extensions.* Differential Geom. Appl. **9**(1998), 173–212.
36. ———, *Reduction by stages and group extensions.* In preparation, 2000.
37. J. E. Marsden and J. Ostrowski, *Symmetries in Motion: Geometric Foundations of Motion Control.* Nonlinear Sci. Today, 1998 (http://link.springer-ny.com).
38. J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry.* 2nd edition, Texts Appl. Math. **17**, Springer-Verlag, 1999.
39. ———, *Reduction of Poisson manifolds.* Lett. Math. Phys. **11**(1986), 161–170.
40. J. E. Marsden, T. S. Ratiu and J. Scheurle, *Reduction theory and the Lagrange-Routh equations.* J. Math. Phys. (6) **41**(2000), to appear.
41. J. E. Marsden, T. S. Ratiu and A. Weinstein, *Semi-direct products and reduction in mechanics.* Trans. Amer. Math. Soc. **281**(1984), 147–177.
42. ———, *Reduction and Hamiltonian structures on duals of semidirect product Lie Algebras.* Contemp. Math. **28**(1984), 55–100.
43. J. E. Marsden and J. Scheurle, *Lagrangian reduction and the double spherical pendulum.* Z. Angew. Math. Phys. **44**(1993), 17–43.
44. ———, *The reduced Euler-Lagrange equations.* Fields Institute Commun. **1**(1993), 139–164.
45. J. E. Marsden and A. Weinstein, *Reduction of symplectic manifolds with symmetry.* Rep. Math. Phys. **5**(1974), 121–130.
46. ———, *The Hamiltonian structure of the Maxwell-Vlasov equations.* Phys. D **4**(1982), 394–406.
47. ———, *Coadjoint orbits, vortices and Clebsch variables for incompressible fluids.* Phys. D **7**(1983), 305–323.
48. J. E. Marsden, A. Weinstein, T. S. Ratiu, R. Schmid, and R. G. Spencer, *Hamiltonian systems with symmetry, coadjoint orbits and plasma physics.* In: Proc. IUTAM-ISIMM Symposium on Modern Developments in Analytical Mechanics (Torino, 1982), Atti della Acad. della Sc. di Torino **117**(1982), 289–340.
49. K. R. Meyer, *Symmetries and integrals in mechanics.* In: Dynamical Systems (ed. M. Peixoto), Academic Press, 1973, 259–273.
50. R. Montgomery, *Canonical formulations of a particle in a Yang-Mills field.* Lett. Math. Phys. **8**(1984), 59–67.
51. ———, *The Bundle Picture in Mechanics.* Ph.D. Thesis, Berkeley, 1986.
52. ———, *The connection whose holonomy is the classical adiabatic angles of Hannay and Berry and its generalization to the non-integrable case.* Comm. Math. Phys. **120**(1988), 269–294.
53. ———, *Isoholonomic problems and some applications.* Comm. Math Phys. **128**(1990), 565–592.
54. ———, *Optimal Control of Deformable Bodies and Its Relation to Gauge Theory.* In: The Geometry of Hamiltonian Systems (ed. T. Ratiu), Math. Sci. Res. Inst. Publ. **22**, Springer-Verlag, 1991.
55. ———, *How much does a rigid body rotate? A Berry's phase from the 18^{th} century.* Amer. J. Phys. **59**(1991), 394–398.

56. ———, *Gauge theory of the falling cat.* Fields Inst. Commun. **1**(1993), 193–218.
57. R. Montgomery, J. E. Marsden, and T. S. Ratiu, *Gauged Lie-Poisson structures.* Contemp. Math. **28**(1984), 101–114.
58. J.-P. Ortega and T. S. Ratiu, *Lectures on Singular Reduction.* To appear, 1999.
59. M. Pedroni, *Equivalence of the Drinfeld-Sokolov reduction to a bi-Hamiltonian reduction.* Lett. Math. Phys. **35**(1995), 291–302.
60. H. Poincaré, *Sur une forme nouvelle des équations de la méchanique.* C. R. Acad. Sci. **132**(1901), 369–371.
61. ———, *Sur la precession des corps deformables.* Bull. Astron. **27**(1910), 321–356.
62. T. S. Ratiu, *The Euler-Poisson equations and integrability.* Thesis, University of California at Berkeley, 1980.
63. ———, *The motion of the free n-dimensional rigid body.* Indiana Univ. Math. J. **29**(1980), 609–629.
64. ———, *Involution theorems.* In: Geometric Methods in Mathematical Physics (eds. G. Kaiser and J. Marsden), Springer Lecture Notes in Math. **775**(1980), 219-257.
65. ———, *Euler-Poisson equations on Lie algebras and the N-dimensional heavy rigid body.* Proc. Nat. Acad. Sci. USA **78**(1981), 1327–1328.
66. ———, *Euler-Poisson equations on Lie algebras and the N-dimensional heavy rigid body.* Amer. J. Math. **104**(1982), 409–448, 1337.
67. E. J. Routh, *Treatise on the Dynamics of a System of Rigid Bodies.* MacMillan, London, 1860.
68. ———, *Stability of a given state of motion.* 1877; reprinted in Stability of Motion (ed. A. T. Fuller), Halsted Press, New York, 1975.
69. ———, *Advanced Rigid Dynamics.* MacMillian and Co., London, 1884.
70. W. J. Satzer, *Canonical reduction of mechanical systems invariant under abelian group actions with an application to celestial mechanics.* Indiana Univ. Math. J. **26**(1977), 951–976.
71. J. C. Simo, D. R. Lewis and J. E. Marsden, *Stability of relative equilibria I: The reduced energy momentum method.* Arch. Rat. Mech. Anal. **115**(1991), 15–59.
72. R. Sjamaar and E. Lerman, *Stratified symplectic spaces and reduction.* Ann. of Math. **134**(1991), 375–422.
73. S. Smale, *Topology and Mechanics.* Invent. Math. **10**(1970), 305–331, **11**, 45–64.
74. J. M. Souriau, *Structure des Systemes Dynamiques.* Dunod, Paris, 1970.
75. S. Sternberg, *Minimal coupling and the symplectic mechanics of a classical particle in the presence of a Yang-Mills field.* Proc. Nat. Acad. Sci. **74**(1977), 5253–5254.
76. P. Vanhaecke, *Integrable Systems in the Realm of Algebraic Geometry.* Lecture Notes in Math. **1638**, Springer, 1996.
77. A. Weinstein, *A universal phase space for particles in Yang-Mills fields.* Lett. Math. Phys. **2**(1978), 417–420.
78. ———, *The local structure of Poisson manifolds.* J. Differential Geom. **18**(1983), 523–557
79. N. Zaalani, *Phase space reduction and Poisson structure.* J. Math. Phys. **40**(1999), 3431–3438.

Control and Dynamical Systems
California Institute of Technology 107-81
Pasadena, CA 91125
USA
email: marsden@cds.caltech.edu

Department of Mathematics
University of California
Berkeley, CA 94720
email: perl@math.berkeley.edu

C. R. Math. Rep. Acad. Sci. Canada Vol. 22 (2), 2000 pp. 55–60

CHARACTERISTIC POLYNOMIALS OF ABELIAN VARIETIES OVER $\mathbb{F}_p$

CHANTAL DAVID

Presented by M. Ram Murty, FRSC

ABSTRACT. Let $P_A(x)$ be the characteristic polynomial of an abelian variety A defined over the finite field $\mathbf{F}_p$. As the roots of $P_A(x)$ are Weil numbers, the triangle inequality gives upper bounds on the coefficients of $P_A(x)$. We show in this paper that there are also lower bounds on the coefficients of a polynomial $P(x)$ of the correct type which implies that $P(x)$ is the characteristic polynomial of an abelian variety over $\mathbf{F}_p$.

RÉSUMÉ. Soit $P_A(x)$ le polynôme caractéristique d'une variété abélienne A sur le corps fini $\mathbf{F}_p$. Comme les racines de $P_A(x)$ sont des nombres de Weil, l'inequalité du triangle donne des bornes supérieures sur les coefficients de $P_A(x)$. Nous montrons dans cet article qu'il y a aussi des bornes inférieures sur les coefficients d'un polynôme $P(x)$ qui impliquent que $P(x)$ est le polynôme caractéristique d'une variété abélienne sur le corps fini $\mathbf{F}_p$.

1. Introduction. Let A be an abelian variety of dimension g over the finite field $\mathbb{F}_p$, and let ϕ be the Frobenius endomorphism of A over $\mathbb{F}_p$. Let $P_A(x)$ be the characteristic polynomial of ϕ acting on the Tate module $T_\ell(A)$ for any $\ell \neq p$. Then, $P_A(x)$ is a monic polynomial of degree $2g$ with coefficients in $\mathbb{Z}$. It was shown by Tate [4] that $P_A(x)$ characterizes A up to $\mathbb{F}_p$-isogeny, *i.e.*, that two abelian varieties A and B over $\mathbb{F}_p$ are $\mathbb{F}_p$-isogenous if and only if $P_A(x) = P_B(x)$.

We write

$$P_A(x) = \prod_{i=1,\ldots,2g} (x - \pi_i).$$

By the Weil's Riemann hypothesis, $|\pi_i| = \sqrt{p}$ for $i = 1, \ldots, 2g$. We then call a Weil number over $\mathbb{F}_p$ an algebraic integer π such that $|\sigma(\pi)| = \sqrt{p}$ for all conjugates of π. Also, the roots of $P_A(x)$ come in pairs $(\pi, p/\pi)$, and $P_A(x)$ writes as

$$(1)\quad P_A(x) = x^{2g} + a_1x^{2g-1} + \cdots + a_gx^g + pa_{g-1}x^{g+1} + \cdots + p^{g-1}a_1x + p^g.$$

We also say that the polynomial $P_A(x)$ is the isogeny class of A.

By the triangle inequality,

$$(2)\qquad |a_n| \leq \binom{2g}{n} p^{n/2} \quad \text{for } n = 1, \ldots, g.$$

Received by the editors February 5, 1999.
AMS subject classification: 14G10.

If A is an elliptic curve, *i.e.*, $g = 1$, then $a_1 = a_p(E)$ and the triangle inequality is the Hasse inequality

$$|a_p(E)| \leq 2\sqrt{p}.$$

Moreover, the triangle inequality is effective in this case, as it was shown by Deuring [1] that for any integer $|a| \leq 2\sqrt{p}$, there is an elliptic curve $E/\mathbb{F}_p$ such that $P_E(x) = x^2 - ax + p$. This is not true for $g > 1$. For example, let $P(x) = x^4 + 3px^2 + p^2$. This is a polynomial of the form (1) satisfying (2), but the roots of $P(x)$ are not Weil numbers over $\mathbb{F}_p$, and then $P(x)$ is not the isogeny class of any abelian variety over $\mathbb{F}_p$.

We show in this paper that there exists constants such that the triangle inequality is effective for all dimensions g. This result also gives a lower bound for the number of isogeny classes of abelian varieties of dimension g over $\mathbb{F}_p$.

THEOREM 1.1. *Fix any integer $g \geq 1$. For any integers $a_1, \dots, a_g$ and any prime p, let $P(x)$ be the polynomial*

$$P(x) = x^{2g} + a_1x^{2g-1} + \cdots + a_gx^g + pa_{g-1}x^{g+1} + \cdots + p^{g-1}a_1x + p^g.$$

Then, there exists positive constants $C_g(n)$ (not depending on p) such

$$|a_n| \leq C_g(n)p^{n/2} \quad \textit{for } n = 1, \dots, g$$

implies that $P(x)$ is the isogeny class of an abelian variety over $\mathbb{F}_p$.

COROLLARY 1.2. *Let g be a positive integer. For any prime p, let $I_g(p)$ be the number of $\mathbb{F}_p$-isogeny classes of abelian varieties of dimension g over $\mathbb{F}_p$. Then, there exists positive constants A_g and B_g depending only on g such that*

$$A_g p^{g(g+1)/4} \leq I_g(p) \leq B_g p^{g(g+1)/4}.$$

PROOF. The lower bound is from Theorem 1.1, and the upper bound from the triangle inequality. ∎

2. Honda's Theorem.

THEOREM 2.1 (HONDA [2]). *Let $q = p^a$ for some positive integer a, and let π be a Weil number over $\mathbb{F}_q$ with minimal polynomial $m(x)$. Then, there exists an elementary abelian variety A over $\mathbb{F}_q$ such that $P_A(x) = m(x)^e$, where e is the least common denominator of the invariants i_v for each place v above p in $E = \mathrm{End}_{\mathbb{F}_q}(A) \otimes \mathbb{Q}$ and*

$$i_v = \begin{cases} \frac{1}{2} & \textit{if } v \textit{ is real;} \\ \frac{f_v \operatorname{ord}_v(\pi)}{a} & \textit{otherwise} \end{cases}$$

where f_v is the residual degree at v.

This version of Honda's Theorem is from [5, Theorem 1]. Then, over $\mathbb{F}_p$, the correspondence between Weil numbers and abelian varieties can be stated as a correspondence between polynomials and abelian varieties. More precisely, we have the following corollary.

COROLLARY 2.2. *Let $P(x) \in \mathbb{Z}[x]$ a polynomial of degree $2g$ of the form*

$$P(x) = x^{2g} + a_1 x^{2g-1} + \cdots + a_g x^g + p a_{g-1} x^{g+1} + \cdots + p^{g-1} a_1 x + p^g$$

such that the roots of $P(x)$ are Weil numbers over $\mathbb{F}_p$. Then, $P(x) = P_A(x)$ for some abelian variety of dimension g over $\mathbb{F}_p$.

PROOF. Write $P(x) = (x^2 - p)^{2n_0} P_1(x)^{n_1} \cdots P_r(x)^{n_r}$ where $n_0 \in \mathbb{N}$ and for $i = 1, \dots, r$, $P_i(x)$ is an irreducible polynomial with no real root, and n_i a positive integer. By Honda's Theorem, there are elementary abelian varieties $A_1, \dots, A_r$ over $\mathbb{F}_p$ such that $P_{A_i}(x) = P_i(x)^{e_i}$ for $i = 1, \dots, r$. As $a = 1$ and $P_i(x)$ has no real roots, $e_i = 1$ for $i = 1, \dots, r$. Similarly, there is an abelian variety A_0 over $\mathbb{F}_p$ such that $P_{A_0}(x) = (x^2 - p)^2$. Then, $P(x) = P_A(x)$ for

$$A = A_0^{n_0} \times A_1^{n_1} \times \cdots \times A_r^{n_r}. \qquad \blacksquare$$

It then suffices to prove the following proposition to obtain Theorem 1.1.

PROPOSITION 2.3. *Fix any integer $g \geq 1$, and let $a_1, \dots, a_g$ be any integers. For any prime p, let $P(x)$ be the polynomial*

$$P(x) = x^{2g} + a_1 x^{2g-1} + \cdots + a_g x^g + p a_{g-1} x^{g+1} + \cdots + p^{g-1} a_1 x + p^g.$$

There exists constants $C_g(n)$, $n = 1, \dots, g$, (not depending on p) such that

$$|a_n| \leq C_g(n) p^{n/2} \quad \text{for } n = 1, \dots, g \tag{3}$$

implies that the roots of $P(x)$ are Weil numbers for $\mathbb{F}_p$.

3. Proof of Proposition 2.3. We number the roots $\pi_1, \dots, \pi_{2g}$ of $P(x)$ such that $\pi_{g+i} = p/\pi_i$ for $i = 1, \dots, g$.

Let $Q(x)$ be the monic polynomial

$$Q(x) = \prod_{i=1,\dots,g} (x - (\pi_i + p/\pi_i)) = \prod_{i=1,\dots,g} (x - \beta_i).$$

LEMMA 3.1. *Let $P(x)$ and $Q(x)$ be the polynomials defined above. Then, $|\pi_i| = \sqrt{p}$ for $i = 1, \dots, 2g$ if and only if $Q(x)$ is totally real, and $|\beta_i| \leq 2\sqrt{p}$ for $i = 1, \dots, g$.*

PROOF. The proof is a standard observation, and can be found for example in [6, p. 528]. If $|\pi_i| = \sqrt{p}$ for $i = 1, \dots, 2g$, then clearly $|\beta_i| \le 2\sqrt{p}$ for $i = 1, \dots, g$. Also, as $\pi_i = p/\pi_i$ if and only if $\pi_i = \pm\sqrt{p}$, it follows that $\beta_i = \pi_i + p/\pi_i$ is either the real part of the complex number π_i or $\pm 2\sqrt{p}$. Conversely, assume that all β_i are real and that $|\beta_i| \le 2\sqrt{p}$. Then, π_i and p/π_i are the roots of

$$x^2 - \beta_i x + p,$$

which has either complex conjugate roots, *i.e.*, $|\pi| = |p/\pi| = \sqrt{p}$, or a double real root, *i.e.*, $\pi = p/\pi = \pm\sqrt{p}$. ∎

It then suffices to prove that (3) implies that $Q(x)$ is totally real with roots β_i such that $|\beta_i| \le 2\sqrt{p}$ for $i = 1, \dots, g$. We write

$$Q(x) = \prod_{i=1,\dots,g} (x - \beta_i) = x^g + b_1 x^{g-1} + \cdots + b_{g-1} x + b_g.$$

We first need to express the coefficients $b_1, \dots, b_n$ of $Q(x)$ in term of $a_1, \dots, a_n$. We also define $a_0 = b_0 = 1$.

LEMMA 3.2. *For* $n = 1, \dots, g$,

$$a_n = b_n + \sum_{i=1}^{[n/2]} \binom{g-(n-2i)}{i} p^i b_{n-2i}.$$

PROOF. Let

$$A_n = (-1)^n a_n = \sum_{1 \le i_1 < \cdots < i_n \le 2g} \pi_{i_1} \cdots \pi_{i_n},$$

and

$$\begin{aligned} B_n &= (-1)^n b_n = \sum_{1 \le i_1 < \cdots < i_n \le g} (\pi_{i_1} + p/\pi_{i_1}) \cdots (\pi_{i_n} + p/\pi_{i_n}) \\ &= \sum_{1 \le i_1 < \cdots < i_n \le 2g}^{(0)} \pi_{i_1} \cdots \pi_{i_n} \end{aligned}$$

where $\sum\limits_{1 \le i_1 < \cdots < i_n \le 2g}^{(r)}$ means that exactly r pairs $(i, g+i)$ occurs among $i_1, \dots, i_n$. Then,

$$\begin{aligned} A_n &= B_n + \sum_{1 \le i_1 < \cdots < i_n \le 2g}^{(1)} \pi_{i_1} \cdots \pi_{i_n} + \sum_{1 \le i_1 < \cdots < i_n \le 2g}^{(2)} \pi_{i_1} \dots \pi_{i_n} + \cdots \\ &= B_n + \sum_{i=1}^{[n/2]} \binom{g-(n-2i)}{i} p^i B_{n-2i}. \end{aligned}$$

This proves the result. ∎

LEMMA 3.3. *For* $n = 1, \dots, g$,

$$b_n = \sum_{i=0}^{[n/2]} d(n,i) p^i a_{n-2i}$$

where $d(n,i)$ *are integers independant of* p.

PROOF. Use Lemma 3.2 and induction on n. ∎

To show Proposition 2.3, we show that the conditions (3) on the coefficients of $P(x)$ imply that the associated polynomial $Q(x)$ is "close enough" to a given polynomial $Q_0(x)$ which is totally real. Consider the polynomial

$$P_0(x) = x^{2g} + p^g,$$

i.e., $a_i = 0$ for $i = 1, \dots, n$. The roots of $P_0(x)$ are

$$\pi_k = \zeta_{4g}^{2k-1} \sqrt{p} \quad \text{for } k = 1, \dots, 2g,$$

where $\zeta_{4g} = \exp(\frac{2\pi i}{4g})$. Then, by Honda's theorem, $P_0(x)$ is the isogeny class of some abelian variety A_0 over $\mathbb{F}_p$. (A_0 is a supersingular abelian variety, *i.e.*, it becomes isogenous to a product of supersingular elliptic curves over some extension of $\mathbb{F}_p$.)

Then

$$Q_0(x) = \prod_{k=1,\dots,g} (x - (\pi_k + p/\pi_k))$$

have roots

$$\beta_k = (\pi_k + p/\pi_k) = 2\sqrt{p}\cos\left(\frac{2\pi(2k-1)}{4g}\right) \quad \text{for } k = 1, \dots, g.$$

For $k = 1, \dots, g-1$, let α_k be non-zero real numbers such that $\beta_{k+1} < \alpha_k < \beta_k$, and let $\alpha_0 = 2\sqrt{p}$ and $\alpha_g = -2\sqrt{p}$. For $k = 0, \dots, g$, we also define γ_k by $\alpha_k = \gamma_k \sqrt{p}$. Then, $Q_0(\alpha_k)$ and $Q_0(\alpha_{k-1})$ have opposite signs for $k = 1, \dots, g$, and if

$$|Q(\alpha_k) - Q_0(\alpha_k)| < |Q_0(\alpha_k)| \quad \text{for } k = 0, \dots, g, \tag{4}$$

then $Q(x)$ also has g real roots of absolute value $\leq 2\sqrt{p}$.

We then have to prove that (3) implies (4). Using Lemma 3.3, we get

$$\begin{aligned} |Q_0(\alpha_k)| &= \Big| \sum_{\substack{n=0 \\ n \text{ even}}}^{g} d(n, n/2) p^{n/2} \gamma_k^{g-n} \sqrt{p}^{g-n} \Big| \\ &= p^{g/2} \Big| \sum_{\substack{n=0 \\ n \text{ even}}}^{g} d(n, n/2) \gamma_k^{g-n} \Big| = p^{g/2} K(g,k), \end{aligned} \tag{5}$$

where $K(g,k)$ is a non-zero constant independant of p. Similarly, if

$$|a_n| \le C_g(n)p^{n/2} \quad \text{for } n = 1, \dots, g,$$

then

$$\begin{aligned}
|Q(\alpha_k) - Q_0(\alpha_k)| &= \Big|\sum_{n=0}^{g}\Big(\sum_{\substack{i=0\\2i\neq n}}^{[n/2]} d(n,i)p^i a_{n-2i}\Big)\gamma_k^{g-n}\sqrt{p}^{g-n}\Big| \\
&\le p^{g/2}\sum_{n=0}^{g}\Big(\sum_{\substack{i=0\\2i\neq n}}^{[n/2]} |d(n,i)|C_g(n-2i)\Big)|\gamma_k|^{g-n}.
\end{aligned} \tag{6}$$

Comparing (5) and (6), it suffices to choose constants $C_g(n)$ such that

$$\sum_{n=0}^{g}\Big(\sum_{\substack{i=0\\2i\neq n}}^{[n/2]} |d(n,i)|C_g(n-2i)\Big)|\gamma_k|^{g-n} < K(g,k) \quad \text{for } k = 0, \dots, g.$$

Let $N(k,g)$ be the number of terms on the sum above. Then, it suffices to choose constants $C_g(n)$ such that for each i, n, k

$$|d(n,i)|C_g(n-2i)|\gamma_k|^{g-n} < \frac{K(g,k)}{N(k,g)}.$$

This completes the proof of Proposition 2.3. ∎

ACKNOWLEDGEMENTS. The author would like to thank Eyal Goren and Hershy Kisilevsky for many helpful discussions, and Francesco Pappalardi for a careful reading of the manuscript.

REFERENCES

1. M. Deuring, *Die Typen de Multiplikatorenringe elliptischer Functionenkörper*. Abh. Math. Sem. Hamburg **14**(1941), 197–272.
2. T. Honda, *Isogeny classes of abelian varieties over finite fields*. J. Math. Soc. Japan **20**(1968), 83–95.
3. J. S. Milne and W. C. Waterhouse, *Abelian varieties over finite fields*. Proc. Sympos. Pure Math. **20**(1971), 53–64.
4. J. Tate, *Endomorphisms of abelian varieties over finite fields*. Invent. Math. **2**(1966), 134–144.
5. ———, *Classes d'isogénie des variétés abéliennes sur un corps fini*. Séminaire Bourbaki, Lecture Notes in Math. **179**, 1971.
6. W. C. Waterhouse, *Abelian varieties over finite fields*. Ann. Sci. École Norm. Sup. **2**(1969), 521–560.

Concordia University
Department of Mathematics
1455 de Maisonneuve Blvd. West
Montréal, Quebec H3G 1M8
email: chantal@cicma.concordia.ca

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 61–65

GENERALIZATION OF LEHMER'S CONGRUENCES FOR BERNOULLI NUMBERS

TAKASHI AGOH

Presented by M. Ram Murty, FRSC

ABSTRACT. Let B_m be the m-th Bernoulli number in the even suffix notation and let $\beta_m = B_m/m$ $(m \neq 0)$. In 1938, E. Lehmer discovered some important congruences involving β_m which are very useful for irregularity testing of primes, and other purposes. In this paper we investigate Lehmer's congruences and derive certain generalizations of them for more general moduli.

RÉSUMÉ. Soit B_m le m-ième nombre de Bernoulli dans les notation de suffixe pair et soit $\beta_m = B_m/m$ $(m \neq 0)$. E. Lehmer dans l'an 1938 decouvrit des congruences importantes concernant les β_m qui son très utiles pour vérifier l'irrégularité de primes et d'autre buts. Dans ce papier nous recherchons les congruences de Lehmer et déduiyons certain généralisations d'eux avec les modules plus généraux.

1. Introduction. Let B_m be the m-th Bernoulli number in the even suffix notation defined by the recurrence

$$B_m = -\frac{1}{m+1}\sum_{i=0}^{m-1}\binom{m+1}{i}B_i, \quad B_0 = 1.$$

Writing B_m for an even $m \geq 2$ as $B_m = N_m/D_m$ $(D_m > 0)$ in lowest terms, we see from the von Staudt-Clausen Theorem that $\mathrm{ord}_p(D_m) = 1$ if $p-1|m$ for a prime p, and $\mathrm{ord}_p(D_m) = 0$ otherwise.

On the other hand, the well-known Euler-MacLaurin summation formula can be stated as

$$S_m(n) = \sum_{j=1}^{m+1}\frac{1}{j}\binom{m}{j-1}B_{m+1-j}n^j, \tag{1.1}$$

where $m, n \geq 1$ and $S_m(n) = 1^m + 2^m + \cdots + (n-1)^m$. Also, it is known that if a is a positive integer with $(a, n) = 1$, then

$$(a^m - 1)S_m(n) = \sum_{i=1}^{m}(-1)^{i-1}n^i\binom{m}{i}\sum_{j=1}^{n-1}(aj)^{m-i}\left[\frac{aj}{n}\right]^i, \tag{1.2}$$

Received by the editors March 3, 1999.

The author was supported in part by a grant of the Ministry of Education, Science and Culture of Japan.

AMS subject classification: 11B68.

where $[\theta]$ is the greatest integer $\le \theta$ for a real θ.

We do not comment on the above two formulas, since these are very familiar in the theory of Bernoulli numbers (see, *e.g.*, [UH]).

Let $\beta_m = B_m/m$ for even $m \ge 2$ and let

$$\begin{aligned} Q_2(X) &= 2^X - 1,\\ Q_3(X) &= \frac{1}{2}(3^X - 1),\\ Q_4(X) &= \frac{1}{2}(2^X - 1)(2^{X-1} + 1),\\ Q_6(X) &= \frac{1}{2}(6^{X-1} + 3^{X-1} + 2^{X-1} - 1). \end{aligned}$$

In 1938, E. Lehmer [L] proved the following theorem which includes very important and useful congruences for irregularity testing of primes, and other purposes.

THEOREM 1.1. *Let p be an odd prime and $m \ge 2$ be even. If $p-1 \nmid m-2$, then*

$$Q_k(m)\beta_m \equiv \sum_{0<i<p/k} (p - ik)^{m-1} \pmod{p^2} \quad (k = 2, 3, 4, 6),$$

provided that $p \ge 7$ for $k = 6$.

We note that a very simple and elegant p-adic proof for Lehmer's theorem was given by Johnson [J].

In this paper we shall extend the congruences in the above theorem to a more general situation and prove the following

THEOREM 1.2. *Let $n \ge 1$ be odd, $m \ge 2$ be even and $\eta = n\prod_{p|n} p^{\mathrm{ord}_p(m)}$ with the product taken over all primes p with $p \mid n$. If $p-1 \nmid m-2$ for each prime divisor p of n, then*

$$Q_k(m)\beta_m \equiv \sum_{0<j<\eta/k} (\eta - jk)^{m-1} \pmod{n^2} \quad (k = 2, 3, 4, 6), \tag{1.3}$$

provided that each prime divisor p of n is ≥ 7 for $k = 6$.

2. Proof of Theorem 1.2. In this section we shall give the proof of Theorem 1.2 after discussing certain consequences of formulas (1.1) and (1.2).

Assume that $m \ge 2$ is even and $p-1 \nmid m-2$ for all prime divisors p of n. Then we have $m \ne 2$ (hence $m \ge 4$) and $B_{m-2} \in \mathbb{Z}_p$ (the ring of all rational numbers which are p-integral) by the von Staudt-Clausen Theorem. Also it is clear that $2 \nmid n$ and $3 \nmid n$, hence any prime divisor p of n is ≥ 5. Recall now (1.1) and consider each term on the right-hand side of this formula. For simplicity, we put

$$M_j = \frac{1}{j}\binom{m}{j-1} B_{m+1-j} n^j \quad \text{for } j = 1, 2, \ldots, m+1.$$

If j is an even integer with $m-2 \geq j \geq 2$, then $M_j = 0$. So we consider here only the terms M_j for $j = 2i+1$ $(i = 0, 1, \ldots, m/2)$ and $j = m$. Let p be any prime divisor of n. If $m > 2i \geq 4$, then $\operatorname{ord}_p(D_{m-2i}) \in \{0, 1\}$, and to estimate $\operatorname{ord}_p(2i+1)$ we note that $p^{2i-3} \geq 5^{2i-3} \geq 2i+1$, hence $\operatorname{ord}_p(M_{2i+1}) \geq -(2i-3) - 1 + (2i+1) = 3$. This gives $M_{2i+1} \equiv 0 \pmod{n^3}$ for all $i \geq 2$. Also, since $\operatorname{ord}_p(D_{m-2}) = 0$ by the assumption, and since $m \geq 4$ and $(n, 6) = 1$, we have $M_3 = \frac{1}{3}\binom{m}{2} B_{m-2} n^3 \equiv 0 \pmod{n^3}$ and $M_m = \frac{1}{m}\binom{m}{m-1} B_1 n^m = -\frac{1}{2} n^m \equiv 0 \pmod{n^3}$. Therefore, we may conclude from (1.1) that

$$S_m(n) \equiv B_m n \pmod{n^3}.$$

On the other hand, we obtain from (1.2) that if $(a, n) = 1$, $a \geq 1$, then

$$\begin{aligned}(a^m - 1)&S_m(n)\\ &\equiv mn \sum_{j=1}^{n-1} (aj)^{m-1} \left[\frac{aj}{n}\right] - \frac{m(m-1)}{2} n^2 \sum_{j=1}^{n-1} (aj)^{m-2} \left[\frac{aj}{n}\right]^2 \pmod{n^3}.\end{aligned}$$

Consequently, we have

$$(a^m - 1) B_m \equiv m \sum_{j=1}^{n-1} (aj)^{m-1} \left[\frac{aj}{n}\right] - \frac{m(m-1)}{2} n \sum_{j=1}^{n-1} (aj)^{m-2} \left[\frac{aj}{n}\right]^2 \pmod{n^2}.$$

Since $(a, \eta) = 1$, the above congruence is also true when n is replaced by η, that is

$$(a^m - 1) B_m \equiv m \sum_{j=1}^{\eta-1} (aj)^{m-1} \left[\frac{aj}{\eta}\right] - \frac{m(m-1)}{2} \eta \sum_{j=1}^{\eta-1} (aj)^{m-2} \left[\frac{aj}{\eta}\right]^2 \pmod{\eta^2},$$

which implies

$$(a^m - 1) \beta_m \equiv \sum_{j=1}^{\eta-1} (aj)^{m-1} \left[\frac{aj}{\eta}\right] - \frac{m-1}{2} \eta \sum_{j=1}^{\eta-1} (aj)^{m-2} \left[\frac{aj}{\eta}\right]^2 \pmod{n^2}.$$

Putting for $i = 0, 1, \ldots, a-1$

$$G_i^{(a)} = \left\{ j \in \mathbb{Z} \;\middle|\; \frac{i\eta}{a} < j < \frac{(i+1)\eta}{a} \right\},$$

and

$$U_i^{(a)} = \sum_{j \in G_i^{(a)}} j^{m-1}, \quad V_i^{(a)} = \sum_{j \in G_i^{(a)}} j^{m-2},$$

the last congruence can be expressed as

$$(2.1) \quad (a^m - 1)\beta_m \equiv a^{m-1} \sum_{i=1}^{a-1} i U_i^{(a)} - \frac{m-1}{2} a^{m-2} \eta \sum_{i=1}^{a-1} i^2 V_i^{(a)} \pmod{n^2}.$$

Next, since m is even, it follows that

$$U_i^{(a)} = \sum_{j\in G^{(a)}_{a-1-i}} (\eta - j)^{m-1} \equiv - \sum_{j\in G^{(a)}_{a-1-i}} j^{m-1} + (m-1)\eta \sum_{j\in G^{(a)}_{a-1-i}} j^{m-2}$$

$$\equiv -U^{(a)}_{a-1-i} + (m-1)\eta V^{(a)}_{a-1-i} \pmod{n^2}. \tag{2.2}$$

We also obtain

$$V_i^{(a)} = \sum_{j\in G^{(a)}_{a-1-i}} (\eta - j)^{m-2} \equiv \sum_{j\in G^{(a)}_{a-1-i}} j^{m-2} \equiv V^{(a)}_{a-1-i} \pmod{n}. \tag{2.3}$$

By the von Staudt-Clausen Theorem $B_{m-2} \in \mathbb{Z}_p$ whenever $p-1 \nmid m-2$ for any prime divisor p of n, therefore we see from (1.1)

$$S_{m-2}(\eta) = \sum_{i=0}^{a-1} V_i^{(a)} \equiv 0 \pmod{n}. \tag{2.4}$$

We are now able to prove Theorem 1.2:

PROOF. The proofs of the formulas for $k = 2, 3, 4, 6$ are similar, therefore we give here the proof only for $k = 6$ which is the most annoying case. The other cases $k = 2, 3$ and 4 are left to the reader.

For brevity, we write $U_i = U_i^{(6)}$ and $V_i = V_i^{(6)}$ $(i = 0, 1, \ldots, 5)$ in what follows. By (2.2) and (2.3) we get $U_i \equiv -U_{5-i} + (m-1)\eta V_{5-i} \pmod{n^2}$ and $V_i \equiv V_{5-i} \pmod{n}$ for $i = 3, 4, 5$. Since $(n, 6) = 1$, taking $a = 6$ in (2.1) we obtain

$$\begin{aligned}(6^m - 1)\beta_m &\equiv 6^{m-1}\sum_{i=1}^{5} iU_i - \frac{m-1}{2}6^{m-2}\eta\sum_{i=1}^{5} i^2 V_i \\ &\equiv -6^{m-1}(5U_0 + 3U_1 + U_2) \\ &\qquad + \frac{m-1}{2}6^{m-2}\eta(35V_0 + 31V_1 + 23V_2) \pmod{n^2}.\end{aligned}$$

It is evident from the given condition that $B_{m-2} \in \mathbb{Z}_p$ for each prime divisor p of n, so we get from (2.4) $\sum_{i=0}^{5} V_i \equiv 2(V_0 + V_1 + V_2) \equiv 0 \pmod{n}$, which implies $V_2 \equiv -V_0 - V_1 \pmod{n}$ because n is odd. Consequently, dividing the above congruence by 6^{m-1} we have

$$(6 - 6^{1-m})\beta_m \equiv -(5U_0 + 3U_1 + U_2) + \frac{m-1}{3}\eta(3V_0 + 2V_1) \pmod{n^2}. \tag{2.5}$$

In a similar way, one can show that

$$\begin{aligned}(2 - 2^{1-m})\beta_m &\equiv -(U_0 + U_1 + U_2) + \frac{m-1}{2}\eta(V_0 + V_1 + V_2) \\ &\equiv -(U_0 + U_1 + U_2) \pmod{n^2}, \\ (3 - 3^{1-m})\beta_m &\equiv -2(U_0 + U_1) + \frac{2(m-1)}{3}\eta(V_0 + V_1) \pmod{n^2}.\end{aligned}$$

These are equivalent to (1.3) for $k = 2$ and $k = 3$ (the details are left to the reader). Subtracting these last two congruences from (2.5), we deduce

$$\begin{aligned}(1 + 2^{1-m} + 3^{1-m} - 6^{1-m})\beta_m &= \{(6 - 6^{1-m}) - (2 - 2^{1-m}) - (3 - 3^{1-m})\}\beta_m \\ &\equiv -2U_0 + \frac{m-1}{3}\eta V_0 \pmod{n^2},\end{aligned}$$

that is,

$$\begin{aligned}(6^{m-1} + 3^{m-1} + 2^{m-1} - 1)\beta_m &\equiv 2\left(-6^{m-1}U_0 + (m-1)6^{m-2}\eta V_0\right) \\ &\equiv 2 \sum_{j \in G_0^{(6)}} (\eta - 6j)^{m-1} \pmod{n^2},\end{aligned}$$

which gives (1.3) for $k = 6$ as desired. ∎

References

[J] W. Johnson, *p-adic proofs of congruences for the Bernoulli numbers.* J. Number Theory 7(1975), 251–265.

[L] E. Lehmer, *On congruences involving Bernoulli numbers and the quotients of Fermat and Wilson.* Ann. of Math. **39**(1938), 350–360.

[UH] J. V. Uspensky and M. A. Heaslet, *Elementary Number Theory.* McGraw-Hill, New York-London, 1939.

Department of Mathematics
Science University of Tokyo
Noda, Chiba 278-8510
Japan
email: agoh@ma.noda.sut.ac.jp

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 66–69

SUR UNE CONJECTURE DE WOLFGANG M. RUPPERT

OMAR KIHEL

Présenté par M. Ram Murty, FRSC

ABSTRACT. If K is a quadratic imaginary number field of discriminant D, Ruppert showed that there exists a generator α of K with minimal polynomial $f(x) = ax^2 + bx + c \in \mathbb{Z}[x]$ and a non-effective constant C such that $H(\alpha) \le C\sqrt{|D|}$ where $H(\alpha) = \max\{|a|, |b|, |c|\}$. He conjectures $C = 3.22$. In this paper, we show that $H(\alpha) \le \frac{p}{\sqrt{3}}\sqrt{|D|}$, where p is a prime which split in $\mathbb{Q}(\sqrt{D})$. As a consequence, we deduce Ruppert's conjecture in several cases.

RÉSUMÉ. Si K est un corps quadratique imaginaire de discriminant D, Ruppert a montré qu'il existe un générateur α de K dont le polynôme minimal est $f(x) = ax^2 + bx + c \in \mathbb{Z}[x]$, et une constante non effective C tel que $H(\alpha) \le C\sqrt{|D|}$ où $H(\alpha) = \max\{|a|, |b|, |c|\}$. Il conjecture que $C = 3.22$. Dans ce papier, on montre que $H(\alpha) \le \frac{p}{\sqrt{3}}\sqrt{|D|}$, où p est un premier qui se décompose dans $\mathbb{Q}(\sqrt{D})$. Comme conséquence, on déduit la conjecture de Ruppert dans plusieurs cas.

1. Introduction. Soit K un corps de nombres de degré n sur le corps des rationnels $\mathbb{Q}$ et α un générateur dont le polynôme minimal est $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ avec $a_i \in \mathbb{Z}$ et $\gcd(a_0, a_1, \ldots, a_n) = 1$. Ruppert a défini dans [2] la hauteur de α comme suit:

$$H(\alpha) = \max(|a_0|, \ldots, |a_n|)$$

et a montré la proposition suivante:

PROPOSITION 1. *Pour tout $n \in \mathbb{N}$, il existe $c_n > 0$ tel que si α engendre un corps de nombres K de degré n et de discriminant D_K alors*

$$H(\alpha) \ge c_n |D_K|^{\frac{1}{2n-2}}.$$

Une question naturelle qui se pose maintenant est la suivante: Existe-t-il un générateur α pour lequel $H(\alpha)$ ait un bon majorant en termes d'invariants du corps K? Ruppert dans [2] a formulé cette question ainsi:

Reçu par les éditeurs le March 15, 1999.
Classification (AMS) par sujet : 11R11, 11R29.

QUESTION 1. *Existe-t-il d_n tel que tout corps de nombres K de degré n sur $\mathbb{Q}$ et de discriminant D_K admet un générateur α tel que:*

$$H(\alpha) \leq d_n |D_K|^{\frac{1}{2n-2}}?$$

Dans le cas où K est un corps quadratique imaginaire de discriminant D, Ruppert [2] a démontré le résultat asymptotique suivant :

$$\lim_{D \to -\infty} \frac{H_{min}(\alpha)}{\sqrt{|D|}} = \frac{1}{2}$$

où, $H_{\min}(\alpha)$ désigne la hauteur d'un générateur de K de hauteur minimale. Ce qui implique l'existence d'une constante (non effective) C tel que $H(\alpha) \leq C\sqrt{|D|}$. Ruppert [2, p. 18] conjecture en fait que $C = 3.22$.

Dans ce travail, on montre (Théorème 1) que $H(\alpha) \leq \frac{p}{\sqrt{3}}\sqrt{|D|}$, où p est le plus petit premier qui se décompose dans $\mathbb{Q}(\sqrt{D})$. On déduit ensuite (Corollaire 1) une réponse affirmative à la conjecture de Ruppert dans les cas suivants:

(1) Si $D \equiv 1 \bmod 3$
(2) Si $D \equiv 1$ ou $4 \bmod 5$
(3) Si $D \equiv 1 \bmod 8$.

THÉORÈME 1. *Si K est un corps quadratique imaginaire de discriminant D et p un premier qui se décompose dans K, alors il existe α tel que $K = \mathbb{Q}(\alpha)$ et*

$$H(\alpha) \leq \frac{p}{\sqrt{3}}\sqrt{|D|}.$$

PREUVE. Soit K un corps quadratique imaginaire de discriminant $D < 0$, alors il existe une infinité de triplets $(a, b, c) \in \mathbb{N}^3$ tel que $D = b^2 - 4ac$. Supposons que $a \leq c$. Parmi ces triplets, choisissons celui qui a le plus petit c. Dans le cas où il y en a plusieurs qui ont le même c (c minimal), alors parmi ces triplets choisissons celui qui a le plus petit a et appelons ce triplet (A, B, C). Ainsi on a :

$$B^2 - D = 4AC,$$

avec $A \leq C$. De l'égalité

$$(B - 2A)^2 - D = 4A(C + A - B),$$

on déduit que $C + A - B \geq C$. D'où $A \geq B$. Ce qui nous donne que

$$A^2 - D \geq 4AC \geq 4A^2.$$

Par conséquent,

$$B \leq A \leq \sqrt{\frac{|D|}{3}}.$$

Soit p un premier qui se décompose dans $\mathbb{Q}(\sqrt{D})$.

(1) Si p divise C, alors

$$B^2 - D = 4pA\frac{C}{p}.$$

D'où $pA \geq C$. Par conséquent,

$$C \leq p\sqrt{\frac{|D|}{3}}.$$

(2) Si p ne divise pas C, alors:

(i) Si p divise A, alors p ne divise pas B car p se décompose dans $\mathbb{Q}(\sqrt{D})$. Donc il existe $k_0 \in \{1, 2, \ldots, p-1\}$ tel que p divise $C - k_0B$. Ainsi $\frac{C+k_0^2A-k_0B}{p}$ est un entier. De l'égalité

$$(B - 2k_0A)^2 - D = 4pA\left(\frac{C + k_0^2A - k_0B}{p}\right),$$

on déduit que $pA \geq C$ ou $\frac{C+k_0^2A-k_0B}{p} \geq C$. Ce qui nous donne que $pA \geq C$. Ainsi on a :

$$C \leq p\sqrt{\frac{|D|}{3}}.$$

(ii) Si p ne divise pas A. Soit β une racine du polynôme $x^2 - Bx + AC$, alors $\mathbb{Q}(\beta) = \mathbb{Q}(\sqrt{D})$. Comme p se décompose dans $\mathbb{Q}(\sqrt{D})$, alors il existe un entier $x_0 \in \{1, 2, \ldots, p-1\}$ tel que $x_0^2 - Bx_0 + AC \equiv 0 \bmod p$. Puisque p ne divise pas A, alors il existe $k_0 \in \{1, 2, \ldots, p-1\}$ tel que $x_0 \equiv Ak_0 \bmod p$. Ainsi $\frac{C+k_0^2A-k_0B}{p}$ est un entier. De la même manière que dans (i), on déduit de l'égalité

$$(B - 2k_0A)^2 - D = 4pA\left(\frac{C + k_0^2A - k_0B}{p}\right),$$

que $pA \geq C$. Ce qui nous donne que

$$C \leq p\sqrt{\frac{|D|}{3}}.$$

En définitif, on a:

$$B \leq A \leq \sqrt{\frac{|D|}{3}} \quad \text{et} \quad C \leq p\sqrt{\frac{|D|}{3}}.$$

Comme $D = B^2 - 4AC$ est le discriminant du corps quadratique $\mathbb{Q}(\sqrt{D})$, alors il est facile de vérifier que $\gcd(A, B, C) = 1$.

Soit α racine du polynôme $Ax^2 - Bx + C$, alors $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{D})$ et $H(\alpha) \leq \frac{p}{\sqrt{3}}\sqrt{|D|}$.

COROLLAIRE 1. *(1) Si $D \equiv 1 \bmod 3$ alors $H(\alpha) \leq \frac{3}{\sqrt{3}}\sqrt{|D|}$.*

(2) Si $D \equiv 1$ *ou* 4 *mod* 5 *alors* $H(\alpha) \leq \frac{5}{\sqrt{3}}\sqrt{|D|}$.

(3) Si $D \equiv 1$ *mod* 8 *alors* $H(\alpha) \leq \frac{2}{\sqrt{3}}\sqrt{|D|}$.

(4) Dans les autres cas, on peut trouver une constante effective en utilisant le théorème 1 si on connait un premier p *qui se décompose dans* $\mathbb{Q}(\sqrt{D})$.

COROLLAIRE 2. *En supposant GRH, on a*

$$H(\alpha) \leq \left(\frac{(70\log(|D|))}{3}\right)^{\frac{1}{2}} \sqrt{|D|}.$$

Ce corollaire se déduit du théorème 1 en utilisant le résultat suivant démontré par Oesterlé[1] (voir [3]): Si on suppose GRH, et si p est le plus petit premier qui se décompose dans le corps quadratique imaginaire K de discriminant D alors $p \leq \left((70\log(|D|))\right)^{\frac{1}{2}}$.

Dans le cas des corps quadratiques réels, Ruppert [2] a démontré le théorème suivant:

THÉORÈME 2. *Si* K *est un corps quadratique réel de discriminant* D*, alors il existe* α *tel que* $K = \mathbb{Q}(\alpha)$ *et*

$$H(\alpha) \leq \sqrt{D}.$$

REFERENCES

1. Hua Loo Keng, *Intoduction to Number Theory*. Springer-Verlag, 1982.
2. Wolfgang M. Ruppert, *Small Generators of Number Fields*. Manuscripta Math. **96**(1998), 17-23.
3. J.-P. Serre, *Quelques applications du théorème de densité de Chebotarev*. Inst. Hautes Études Sci. Publ. Math. **54**(1981), 123-201.

Département de mathématiques et de statistique
Cité universitaire, Pavillon Vachon
Université Laval
Ste-Foy, Québec
G1K 7P4
email: kihel@mat.ulaval.ca

[1] Je tiens à remercier le professeur K. Murty de m'avoir communiqué la version exacte du théorème d'Oesterlé.

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 70–76

CLASSICAL BIORTHOGONAL RATIONAL FUNCTIONS ON ELLIPTIC GRIDS

VYACHESLAV SPIRIDONOV AND ALEXEI ZHEDANOV

Presented by Vlastimil Dlab, FRSC

ABSTRACT. We construct a new family of rational functions which are biorthogonal with a discrete weight on the grids described by elliptic Jacobi functions. These rational functions satisfy a three-term recurrence relation and are expressed in terms of an elliptic generalization of the very-well-poised balanced basic hypergeometric series ${}_{10}\varphi_9$.

RÉSUMÉ. On construit une noùvelle famille de fonctions rationnelles, biorthogonales par rapport à un poid discret, défini sur les réseaux décrits par les fonctions elliptiques de Jacobi. Ces fonctions rationnelles satisfont une relation de récurrence à trois termes et elles sont exprimées par une généralisation elliptique de la fonction hypergéometrique de base ${}_{10}\varphi_9$ trés bien équilibrée et balancée.

1. Introduction. Orthogonal polynomials (OP) $P_n(z)$ are known to satisfy the three-term recurrence relation

$$J\, P_n(z) = zP_n(z), \tag{1.1}$$

where J is a Jacobi (*i.e.*, tri-diagonal) matrix acting on the polynomials as follows

$$J\, P_n(z) = a_n\, P_{n+1}(z) + b_n\, P_n(z) + u_n\, P_{n-1}(z). \tag{1.2}$$

Today it is commonly believed that all classes of OP explicitly expressed through basic hypergeometric functions are related to the Askey-Wilson polynomials [1] or their limiting cases. The Askey-Wilson polynomials are determined by the basic hypergeometric function ${}_4\varphi_3$.

In his thesis [11] and in [12] Wilson has constructed a large family of finite-dimensional *biorthogonal rational* functions. Namely, a particular pair of rational functions $R_n(z)$, $T_n(z)$ was presented which satisfy the relation

$$\sum_{s=0}^{N} R_n(z_s)T_m(z_s)w_s = h_n\delta_{nm} \tag{1.3}$$

for the so-called q-quadratic grid [6], [8]

$$z_s = C_1q^s + C_2q^{-s} + C_3. \tag{1.4}$$

Received by the editors June 10, 1999.
AMS subject classification: Primary: 33E05; secondary: 39A10.

These functions $R_n(z)$, $T_n(z)$ are expressed in terms of the very-well-poised balanced basic hypergeometric function ${}_{10}\varphi_9$ (see, *e.g.*, [3] for notation). For other examples of biorthogonal rational functions and further development of the general theory see, *e.g.* [4], [5], [7], [8].

In [5] and [13] it was shown that in general the functions $R_n(z)$ satisfy the three-term recurrence relation

$$J_1 R_n(z) = z\, J_2 R_n(z), \tag{1.5}$$

where $J_{1,2}$ are two arbitrary Jacobi matrices. A similar equation is satisfied by $T_n(z)$. In contrast to (1.1) the equation (1.5) belongs to the class of generalized eigenvalue problems [10]. It is well known that solutions of such equations satisfy biorthogonality relation (instead of the pure orthogonality for eigenfunctions of the ordinary eigenvalue problems). This is why the finite sets of rational functions $R_n(z)$, $T_n(z)$ satisfy the relation (1.3).

In analogy with the Askey-Wilson polynomials situation, it was expected that biorthogonal rational functions considered in [4], [8], [12], which are associated with the grids (1.4), are the most general ones satisfying the three-term recurrence relation (1.5) and admitting an explicit expression for $R_n(z)$, $T_n(z)$ in terms of some series of the hypergeometric type.

In this paper we describe the functions $R_n(z)$, $T_n(z)$ which are biorthogonal on the so-called elliptic grids. These functions are expressed in terms of the "modular hypergeometric functions" introduced in [2]. Thus the class of biorthogonal rational functions satisfying three-term recurrence relation (1.5) and expressed in terms of some generalizations of the hypergeometric series is much wider than it was guessed earlier, see *e.g.* [4]. Complete proofs of the announced results will be published separately (some details can be found in [9]).

2. Elliptic analogues of hypergeometric functions. Let $\theta_1(u)$ be the Jacobi theta function defined as (see, *e.g.* [3]):

$$\begin{aligned}\theta_1(u) &= 2\sum_{n=0}^{\infty}(-1)^n p^{(n+1/2)^2}\sin(2n+1)u \\ &= 2p^{1/4}\sin u\prod_{n=1}^{\infty}(1-2p^{2n}\cos 2u+p^{4n})(1-p^{2n}),\end{aligned} \tag{2.1}$$

where p is a complex parameter, $|p| < 1$. The modular parameter τ is introduced in the standard way $p = e^{\pi i\tau}$.

Following [2], define the "elliptic numbers" (or, simply, e-numbers):

$$[x; h, \tau] = \frac{\theta_1(\pi h x)}{\theta_1(\pi h)}, \tag{2.2}$$

where h is an arbitrary constant. The e-numbers depend on three variables x, h and τ. In what follows the dependence on h, τ will be omitted in the notation, *i.e.*, we assume that $[x] \equiv [x; h, \tau]$. We will use also the notations

$$[x]_0 = 1, \quad [x]_n = [x][x+1]\cdots[x+n-1], \quad [n]! = [1]_n,$$

which are natural elliptic generalizations of the Pochhammer symbol and factorial.

The elliptic analogue of the very-well-poised balanced hypergeometric functions ${}_{r+1}E_r$ is defined by the formal series [2]

$$(2.3) \quad {}_{r+1}E_r(a_1; a_4, a_5, \ldots, a_{r+1}; h, \tau) = \sum_{k=0}^{\infty} \frac{[a_1]_k [a_1+2k]}{[k]!\,[a_1]} \prod_{m=4}^{r+1} \frac{[a_m]_k}{[1+a_1-a_m]_k}$$

with the balancing restriction upon the parameters

$$(2.4) \qquad \frac{r-5}{2} + \frac{r-3}{2} a_1 - \sum_{m=4}^{r+1} a_m = 0.$$

We have chosen notations different from those of [2], in particular, the letter 'E' was taken for an association with the elliptic functions.

For $p \to 0$ the series (2.3) becomes the very-well-poised balanced basic series ${}_{r+1}\varphi_r$ for $q = e^{2\pi i h}$. It should be noted that in general the sum in (2.3) diverges in the same way as most of the infinite q-series do for $|q| = 1$. In order to avoid this problem, we restrict the situation to the case when one of the parameters (say, a_{r+1}) is equal to a negative integer $a_{r+1} = -n$. In this case (2.3) terminates at $k = n + 1$ and becomes a finite sum of elliptic functions. Another essential remark is that there is no coalescence procedure—the limits $a_m \to \infty$ are not well defined and they do not simplify the series. Fixing one of the parameters a_m equal to 1 and some others equal to $(1+a_1)/2$ one can arrive at the simplest possible form of (2.3), but, still, many of simple q-series formulae do not have elliptic generalizations.

An analogue of the Bailey's transformation formula which was proven in [2] has the form:

$$(2.5) \qquad \begin{aligned} &{}_{10}E_9(a_1; a_4, \ldots, -n; h, \tau) \\ &= {}_{10}E_9(b_1; b_4, \ldots, b_{10}; h, \tau) \\ &\quad \times \frac{[a_1+1]_n [b_1+1-a_7]_n [b_1+1-a_8]_n [a_1+1-a_7-a_8]_n}{[b_1+1]_n [a_1+1-a_7]_n [a_1+1-a_8]_n [b_1+1-a_7-a_8]_n}, \end{aligned}$$

where

$$(2.6) \qquad \begin{aligned} b_1 = 2a_1 + 1 - a_4 - a_5 - a_6, &\quad b_4 = b_1 - a_1 + a_4, \\ b_5 = b_1 - a_1 + a_5, &\quad b_6 = b_1 - a_1 + a_6 \end{aligned}$$

and $\{b_7, b_8, b_9, b_{10}\}$ is an arbitrary permutation of the parameters $a_7, a_8, a_9, a_{10} = -n$.

The next two theorems are new and describe generalizations of the contiguous relations for the corresponding basic hypergeometric functions ${}_{10}\varphi_9$ from [4]. In what follows we use for brevity the notations similar to the ones suggested in [4]: $\Phi(a_i\pm)$ denotes the function ${}_{10}E_9$ with the particular parameter a_i replaced by $a_i \pm 1$, whereas other parameters being unchanged. $\Phi_\pm$ denotes the function ${}_{10}E_9(a_1 \pm 2; a_4 \pm 1, \dots, a_{10} \pm 1; h, \tau)$.

THEOREM 1. *Assume that one of the parameters $a_i = -n$, where $i = 4, \dots, 10$. Then the following identity takes place*

$$\Phi(a_9-, a_{10}+) - \Phi = \Phi_+(a_9-)\frac{[a_1+1][a_1+2][a_{10}-a_9+1][a_{10}+a_9-a_1-1]}{[1+a_1-a_9][2+a_1-a_9][a_1-a_{10}][1+a_1-a_{10}]} \prod_{i=4}^{8} \frac{[a_i]}{[1+a_1-a_i]}. \tag{2.7}$$

THEOREM 2. *Under the same assumptions as in the previous theorem, the following identity takes place*

$$\begin{aligned} &\frac{[a_{10}]}{[1+a_1-a_9][2+a_1-a_9]} \prod_{i=4}^{8} [1+a_1-a_i-a_9]\Phi_+(a_9-) \\ &= \frac{[a_9]}{[1+a_1-a_{10}][2+a_1-a_{10}]} \prod_{i=4}^{8} [1+a_1-a_i-a_{10}]\Phi_+(a_{10}-) \\ &+ \frac{[a_{10}-a_9]}{[1+a_1][2+a_1]} \prod_{i=4}^{8} [1+a_1-a_i]\Phi. \end{aligned} \tag{2.8}$$

3. Elliptic rational functions. Let us describe a class of biorthogonal rational functions $R_n(z)$, $T_n(z)$ expressed explicitly in terms of the elliptic analogues of hypergeometric functions ${}_{10}E_9$.

Let $d_1, \dots, d_5, e_1, e_2$ and x_0, x_1, x_2 be arbitrary parameters with the restrictions

$$x_2 = x_0 + x_1, \quad e_1 + e_2 = d_1 + d_2, \quad d_3 = x_2 - N - 1, \quad \sum_{i=1}^{5} d_i = 1 + 2(x_0 + x_2),$$

where N is a fixed positive integer. Assume also that $h \neq m_1 + m_2\tau$, $d_1 + d_2 - 2x_2 \neq (m_1 + m_2\tau)/h - M - 2$, $d_1 + d_2 \neq (m_1 + m_2\tau)/h + M + 2$ and $d_1 + d_2 - 2x_0 \neq (m_1 + m_2\tau)/h + M + 2$ for any integers $m_{1,2} = 0, \pm1, \pm2, \dots$ and $M = 1, 2, \dots$. Define the functions

$$\begin{aligned} R_n(z(u)) = {}_{10}E_9(&1 - x_1; 2 - x_1 + N, 1 + x_0 - d_4, 1 + x_0 - d_5, -n, \\ &1 - x_2 + n, 1 + u + x_0 - e_1, 1 - u + x_0 - e_2; h, \tau) \end{aligned} \tag{3.1}$$

and

$$T_n(z(u)) = {}_{10}E_9(2+x_0-d_1-d_2; 3+x_0-d_1-d_2+N, -x_0+d_4-N, -x_0+d_5-N, -n, 1-x_2+n, 1+u+x_0-e_1, 1-u+x_0-e_2; h, \tau). \tag{3.2}$$

THEOREM 3. *The functions* $R_n(z)$ *and* $T_n(z)$ *are rational functions of the argument*

$$z(u) = \frac{[u][u+e_2-e_1]}{[u+d_2-e_1][u+d_1-e_1]} \tag{3.3}$$

with the simple poles of $R_n(z)$ *located at* $z = \alpha_k$,

$$\alpha_k = \frac{[k-x_2+e_1][k-x_2+e_2]}{[k-x_2+d_1][k-x_2+d_2]}, \quad k = 1, 2, \ldots, n, \tag{3.4}$$

and simple poles of $T_n(z)$ *located at* $z = \beta_k$,

$$\beta_k = \frac{[k-e_1+1][k-e_2+1]}{[k-d_1+1][k-d_2+1]}, \quad k = 1, 2, \ldots, n. \tag{3.5}$$

The functions $R_n(z)$ *and* $T_n(z)$ *are biorthogonal*

$$\sum_{s=0}^{N} R_n(z_s) T_m(z_s) \omega_s = h_n \delta_{nm} \tag{3.6}$$

on the "elliptic grid"

$$z_s = \frac{[s+1+x_0-e_1][s+1+x_0-e_2]}{[s+1+x_0-d_1][s+1+x_0-d_2]}, \quad s = 0, 1, 2, \ldots, N$$

with the weight function

$$\omega_s = \frac{[2x_0+2-d_1-d_2+2s][-N]_s[2x_0+2-d_1-d_2]_s}{[2x_0+2-d_1-d_2][s]!\,[2x_0+3-d_1-d_2+N]_s} \times \frac{[x_0]_s[1+d_4-x_2]_s[1+d_5-x_2]_s[1+x_0+x_2-d_1-d_2]_s}{[2-x_1]_s[3+x_0-d_1-d_2]_s[-N+d_4]_s[-N+d_5]_s} \tag{3.7}$$

and the normalization constants

$$h_n = \kappa \frac{[1-x_2][n]!\,[2-x_2+N]_n}{[1-x_2+2n][-N]_n[1-x_2]_n} \times \frac{[2-x_1]_n[3+x_0-d_1-d_2]_n[1-d_4]_n[1-d_5]_n}{[-1-x_0-x_2+d_1+d_2]_n[1+d_4-x_2]_n[1+d_5-x_2]_n[-x_0]_n}, \tag{3.8}$$

where

$$\kappa = \frac{[2-x_2]_N[x_2-d_4-d_5]_N[1+x_0-d_4]_N[1+x_0-d_5]_N}{[1-d_4]_N[1-d_5]_N[2-x_1]_N[x_0+x_2-d_4-d_5]_N}. \tag{3.9}$$

THEOREM 4. *The rational functions $R_n(z)$ satisfy the three-term recurrence relation*

$$\begin{aligned} &\epsilon_n a_n(z-\alpha_{n+1})\big(R_{n+1}(z)-R_n(z)\big)-\epsilon_{n-1}b_n(z-\beta_{n-1})\big(R_n(z)-R_{n-1}(z)\big) \\ &= c_n(z-z_0)R_n(z), \quad n=0,1,\ldots, \end{aligned} \tag{3.10}$$

where

$$\epsilon_n = \frac{[2+n-x_1][3+n-x_1][x_0-n][x_0-n+1]}{[2-x_1][3-x_1][2-x_2+2n][-x_0-1]}\prod_{i=1}^{5}\frac{[1-x_2+d_i]}{[1+x_0-d_i]},$$

$$a_n = \frac{[n+1-x_2]}{[2+n-x_1][3+n-x_1]}\prod_{i=1}^{5}[n+1-x_2+d_i],$$

$$b_n = \frac{[-n]}{[1+x_0-n][2+x_0-n]}\prod_{i=1}^{5}[d_i-n],$$

$$c_n = \frac{[1-x_2+2n]}{[2-x_1][3-x_1]}\prod_{i=1}^{5}[1-x_2+d_i].$$

REMARK. Since $b_0=0$ it is assumed that the second term in (3.10) vanishes identically for $n=0$.

Finally, we describe a self-duality property which allows one to use the adjective "classical" as far as the derived set of biorthogonal rational functions is concerned.

THEOREM 5. *Consider the following transformation of the parameters*

$$\begin{aligned} &\tilde{x}_0=-1-x_0-x_2+d_1+d_2, \quad \tilde{x}_1=x_1, \quad \tilde{x}_2=-1-2x_0+d_1+d_2, \\ &\tilde{d}_3=-2-N-2x_0+d_1+d_2, \quad \tilde{d}_4=-1-2x_0-x_2+d_1+d_2+d_4, \\ &\tilde{d}_5=-1-2x_0-x_2+d_1+d_2+d_5, \quad \tilde{d}_1+\tilde{d}_2=2d_1+2d_2-1-2x_0-x_2. \end{aligned} \tag{3.11}$$

It is directly verified that

$$\tilde{R}_n(z_s)=R_s(z_n), \quad \tilde{T}_n(z_s)=T_s(z_n), \quad 0\le s,n\le N, \tag{3.12}$$

where $\tilde{R}_n(z_s)$, $\tilde{T}_n(z_s)$ denotes the functions obtained from $R_n(z_s)$, $T_n(z_s)$ after replacing the parameters $d_1,\ldots,x_2$ by $\tilde{d}_1,\ldots,\tilde{x}_2$. Thus the transformation (3.11) is equivalent to the permutation of discrete variables n and s.

ACKNOWLEDGMENTS. This work is supported in part by INTAS grant 96-0700 and RFBR (Russia) grant 97-01-00281. A. Zh. thanks Centre de recherches mathématiques of the Université de Montréal for hospitality.

REFERENCES

1. R. Askey and J. Wilson, *Some Basic Hypergeometric Orthogonal Polynomials that Generalize Jacobi Polynomials*. Mem. Amer. Math. Soc. **54**(1985), 1–55.
2. I. B. Frenkel and V. G. Turaev, *Elliptic solutions of the Yang-Baxter equation and modular hypergeometric functions*. In: The Arnold-Gelfand Mathematical Seminars, Birkhäuser, Boston, MA, 1997, 171–204.
3. G. Gasper and M. Rahman, *Basic Hypergeometric Series*. Cambridge University Press, Cambridge, 1990.
4. D. P. Gupta and D. R. Masson, *Contiguous relations, continued fractions and orthogonality*. Trans. Amer. Math. Soc. **350**(1998), 769–808.
5. M. E. H. Ismail and D. R. Masson, *Generalized orthogonality and continued fractions*. J. Approx. Theory **83**(1995), 1–40.
6. A. F. Nikiforov, S. K. Suslov and V. B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable*. Springer, 1991.
7. M. Rahman, *An integral representation of a* ${}_{10}\varphi_9$ *and continuous biorthogonal* ${}_{10}\varphi_9$ *rational functions*. Canad. J. Math. **38**(1986), 605–618.
8. M. Rahman and S. Suslov, *Classical biorthogonal rational functions*. In: Methods of Approximation Theory in Complex Analysis and Mathematical Physics IV, Lecture Notes in Math. **1550**, Springer-Verlag, Berlin, 1993, 131–146.
9. V. Spiridonov and A. Zhedanov, *Spectral transformation chains and some new biorthogonal rational functions*. Comm. Math. Phys., to appear.
10. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.
11. J. A. Wilson, *Hypergeometric series, recurrence relations and some new orthogonal functions*. Ph.D. thesis, University of Wisconsin, Madison, WI, 1978.
12. ———, *Orthogonal functions from Gram determinants*. SIAM J. Math. Anal. **22**(1991), 1147–1155.
13. A. Zhedanov, *Biorthogonal rational functions and generalized eigenvalue problem*. J. Approx. Theory, to appear.

Bogoliubov Laboratory of Theoretial Physics
JINR
Dubna
Moscow region 141980
Russia
email: *svp@thsun1.jinr.ru*

Donetsk Institute for Physics and Technology
Donetsk 340114
Ukraine
email: *zhedanov@host.dipt.donetsk.ua*
and
Centre de recherches mathématiques
Université de Montréal
C.P. 6128, succ. Centre-Ville
Montréal QC
H3C 3J7
email: *zhedanov@crm.umontreal.ca*

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 77–81

REPRESENTATIONS OF COMPLETELY SOLVABLE LIE ALGEBRAS OVER A RING OF POLYNOMIALS

ALEXANDER GRISHKOV

Presented by Vlastimil Dlab, FRSC

RÉSUMÉ. Si R est un anneau de polynômes et L est une R-algèbre complètement résoluble, c'est-à-dire qu'il existe une chaine d'idéaux $L = L_0 \subset L_1 \subset \cdots \subset L_n \subset 0$, avec L_i/L_{i+1} un R-module libre de dimension 1, alors L admet une représentation par matrices triangulaires sur R.

In this article we show that a completely solvable Lie algebra over a ring of polynomials has a triangulable faithful representation. This fact is important in the theory of analytic diassosiative loops. In 1955 Malcev noted [5] that, every Binary Lie algebra B over the field $\mathbf{R}$ of real numbers, some neighborhood $D \subset B$ of zero can be considered as a local analytic diassociative loop with the multiplication $x \star y = H(x, y) = x + y + [x, y]/2 + \cdots$, where $H(x, y)$ stands for the Campbell-Hausdorf series. Recall that by definition an algebra B is Binary Lie if every two elements of B generate a Lie subalgebra. By the Cartan theorem, for every local analytic Lie group, there exists a corresponding global analytic Lie group. An analogy is not valid for local analytic diassociative loops [1]. The problem of existence of a global analytic dissociative loop for a given Binary Lie algebra is not easy. The following can help in solving the problem.

Let B be a Binary Lie algebra over $\mathbf{R}$ with a finite basis $\{b_1, \ldots, b_n\}$ and $B_K = B \otimes_R K$ be corresponding Binary Lie over the polynomial ring $K = \mathbf{R}[x_1, x_2, \ldots, y_1, y_2, \ldots]$. Denote $X = x_1 b_1 + \cdots + x_n b_n$ and $Y = y_1 b_1 + \cdots + y_n b_n$, and let $B(X, Y)$ be a Lie subalgebra generated by X, Y. By the theorem [2], the algebra B is completely solvable, hence, $B(X, Y)$ is so. By Theorem 1 of the present work $B(X, Y)$ has a faithful triangulable representation π. Therefore, there exists a triangular matrix Z such that $\exp(\pi(X)) \exp(\pi(Y)) = \exp(Z)$. Suppose that $Z = f_1 \pi(b_1) + \cdots + f_n \pi(b_n)$, where $f_1, \ldots, f_n$ are analytic functions on $\mathbf{R}^{2n}$. Then $\mathbf{R}^n$ is a global analytic diassociative loop with a multiplication given by the rule

$$(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = (f_1, \ldots, f_n).$$

Moreover, this loop corresponds to B.

Received by the editors June 10, 1999.
AMS subject classification: 17B05, 17B20, 17B67.

Let R be a polynomial ring, $R = k[x_1, \dots, x_m]$, where k is a field of characteristic 0, let Q be the field of fractions of R. For every R-module V, we denote by $\dim_R V$ the dimension of the Q-module $V \otimes_R Q$. We call an Lie algebra L over R *completely solvable* if L admits some normal series

$$L = L_1 > L_2 > \cdots > L_n = 0,$$

such that $\dim_R(L_i/L_{i+1}) = 1$, $i = 1, \dots, n-1$.

A module V over the Lie algebra L_R is called *triangulable* if V is free R-module with a free basis $\{v_1, \dots, v_p\}$, such that $V_i = \{v_i, \dots, v_p\}_R$ is an L-submodule.

The main result of this section is the following:

THEOREM 1. *Let $R = k[x_1, \dots, x_m]$, L be a completely solvable Lie algebra over R, and let L be free as an R-module. Then L has a faithful triangulable module.*

PROOF. Suppose that L is a splitting Lie algebra, *i.e.*, $L = T \oplus N$, where N is the nilradical of L and T is a torus. Moreover, T and N are free R-module. Let $t_1, \dots, t_q$ be a free basis of the R-module T. First we will prove the statement in this particular case by induction on q.

1. CASE $q = 0$. Let $\mathcal{L} = L \otimes_R Q$, a nilpotent Lie algebra over Q. By the Ado theorem [3], $\mathcal{L}$ has a faithful finite-dimensional representation $\rho: \mathcal{L} \to \mathrm{End}_Q V$ such that all elements $\rho(x)$, $x \in \mathcal{L}$ are nilpotent. Fix a Q-basis in V and generators $n_1, \dots, n_l$ of the R-module $\mathcal{L}$. We can find $a \in R$ such that all coefficients of the matrices $a\rho(n_1), \dots, a\rho(n_l)$ are in R. We denote $D = \mathrm{diag}(1, a, \dots, a^{n-1})$, where $n = \dim_Q V$. If the matrices $\rho(n_1), \dots, \rho(n_l)$ are upper triangular, then the matrices

$$D^{-1}\rho(n_1)D, \dots, D^{-1}\rho(n_l)D$$

have coefficients in R. It means that the Lie R-algebra L has faithful triangable module over R.

2. CASE $q > 0$. Denote $t = t_q$ and $L_0 = T_0 \oplus N$, where $T_0 = \{t_1, \dots, t_{q-1}\}_R$. Now it suffices to prove the following:

LEMMA 1. *Let L be a completely solvable Lie R-algebra and let $L = Rt \oplus L_0$, where L_0 is an ideal of L. Then, for every representation $\rho_0: L_0 \to \mathrm{End}_R(V_0)$, where V_0 is a free R-module, there exists an extension $\rho: L \to \mathrm{End}_R(V)$ such that V is a free R-module and $V = V_0 \oplus V_1$, V_1 is an R-submodule of V, $\rho|_{V_0} = \rho_0$. Moreover, if V_0 is a triangulable L_0-module, then V is triangulable.*

PROOF. Let W be an L-module induced by V_0, $W = \sum_{i=0}^{\infty} \oplus V_0 \otimes t^i$. The lemma will be proved if we construct a unitary polynomial $f(x) \in R[x]$ such that $V_0 f(t) L \cap V_0 = 0$ with all the roots of $f(x)$ are in R. We will construct $f(x)$ by induction on $\dim V_0$. If $\dim V_0 = 1$, we can take $f(x) = x^2$. In the case of $\dim V_0 > 0$, we need the following:

LEMMA 2. *Let $s(x)$ and $p(x)$ be unitary polynomials, $s(x), p(x) \in R[x]$. Then there exists an unitary polynomial $f(x) \in R[x]$ such that, in the ring $R[x,y]$,*

$$f(x+y) = f(y) + s(x)r_1(x,y) + xp(x)r_2(x,y) \tag{1}$$

for some $r_1(x,y), r_2(x,y) \in R[x,y]$. Moreover, if all the roots of $s(x)$ and $p(x)$ are in R then all the roots of $f(x)$ are in R.

PROOF. Let R_1 be an extension of R such that $s(x) = (x-\alpha_1)\cdot \cdots \cdot (x-\alpha_n)$ and $p(x) = (x-\beta_1)\cdot \cdots \cdot (x-\beta_m)$ for some $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \in R_1$. We define

$$f(x) = p^n(x)\prod_{i,j}(x-\alpha_i-\beta_j).$$

Note that $f(x) \in R[x]$ if $s(x), p(x) \in R[x]$. We will prove (1) by induction on n. For $n=1$,

$$\begin{aligned}
&f(x+y) - f(x) \\
&\quad = p(x+y)\prod_j(x+y-\alpha_1-\beta_j) - p(y)\prod_j(y-\alpha_1-\beta_j) \\
&\quad = p(x+y)\prod_j[(x-\alpha_1)+(y-\beta_j)] - p(y)\prod_j(y-\alpha_1-\beta_j) \\
&\quad = p(x+y)\prod_j(y-\beta_j) + (x-\alpha_1)r_1'(x,y) - p(y)\prod_j(y-\alpha_1-\beta_j) \\
&\quad = (x-\alpha_1)r_1'(x,y) + p(y)\Big[\prod_j(x+y-\beta_j) - \prod_j(y-\alpha_1-\beta_j)\Big].
\end{aligned}$$

Since

$$\prod_j(x+y-\beta_j) = xr_3(x,y) + p(y),$$

$$\begin{aligned}
\prod_j(y-\alpha_1-\beta_j) &= \prod_j[(x-\alpha_1)+(y-\beta_j)-x] \\
&= (x-\alpha_1)r_4(x,y) + xr_5(x,y) + p(y),
\end{aligned}$$

we obtain

$$\Big[\prod_j(x+y-\beta_j) - \prod_j(y-\alpha_1-\beta_j)\Big] = x(r_3-r_5) - (x-\alpha_1)r_4.$$

The equality (1) is proved for $n=1$.

Let $n>1$. Denote $s_1(x) = \prod_i^{n-1}(x-\alpha_i)$. By induction,

$$\begin{aligned}
&f(x+y) - f(y) \\
&\quad = p^n(x+y)\prod_{i,j}(x+y-\alpha_i-\beta_j) - f(y)
\end{aligned}$$

$$= \Big\{p^{n-1}(x+y)\prod_{i,j}(x+y-\alpha_i-\beta_j)\Big\}\Big\{p(x+y)\prod_j(x+y-\alpha_n-\beta_j)\Big\}$$
$$-f(y)$$
$$= \Big\{s_1(x)r_1'(x,y)+xp(y)r_2'(x,y)+p^{n-1}(y)\prod_{i<n,j}(y-\alpha_i-\beta_j)\Big\}$$
$$\cdot\Big\{(x-\alpha_n)r_3'(x,y)+xp(y)r_4'(x,y)+p(y)\prod_j(y-\alpha_n-\beta_j)\Big\}$$
$$-p^n(y)\prod_{i,j}(y-\alpha_i-\beta_j) = s(x)r_1'(x,y)r_3'(x,y)+xp(y)r_2(x,y). \quad \blacksquare$$

Now we return to the proof of Lemma 1. As L_0 is triangulable L_0-module, then there exists an L_0-submodule $V_1 < V_0$ such that $V_0 = V_1 \oplus Rv$, $v \in V_0$. By induction, there exists a polynomial $p(x)$ such that the induced L-module $W_1 = \sum_{i=0}^{\infty} \oplus V_1 t^i$ contains a submodule $V_2 \supset V_1 p(t)$ and $V_2 \cap V_1 = 0$. Moreover, W_1/V_2 is a triangulable L-module. Denote by $s(x)$ a minimal polynomial such that $Ls\big(ad(t)\big) = 0$. By induction, all the roots of $p(x)$ are in R and all the roots of $s(x)$ are in R, since L is completely solvable. By Lemma 2, to given $s(x)$ and $p(x)$, we can associate a polynomial $f(x)$ such that equality (1) holds. We will prove that $W_0 = V_2 + Wf(t)$ is an L-submodule what we need. As $W_0 t \subseteq W_0$, we have to prove that $Wf(t)a \subseteq W_0$ for any $a \in L_0$. Denote $a^i = [a, \underbrace{t, \dots, t}_{i}]$, $v \in W$. From [3, p. 50], it follows that

$$vf(t)a = \sum_{i=0}^{\infty} va^i f^{(i)}(t)/i!\,. \tag{2}$$

By Lemma 2, in the commutative ring $R[x,y]$ we get

$$\begin{aligned}\sum_{i=0}^{\infty} x^i f^{(i)}(y)/i! &= f(x+y)\\ &= f(y)+s(x)r_1(x,y)+xp(y)r_2(x,y)\\ &= f(y)+s(x)\sum_{i=0}^{k_1} r_{1i}(x)y^i+\sum_{i=1}^{k_2} r_{2i}(x)p(y)y^i,\end{aligned} \tag{3}$$

where

$$r_1(x,y) = \sum_{i=0}^{k_1} r_{1i}(x)y^i, r_2(x,y) = \sum_{i=0}^{k_2} r_{2i}(x)y^i.$$

Denote $s_i(x) = s(x)r_{1i}(x)$. From (4), we derive

$$\sum_{i=0}^{\infty} x^i f^{(i)}(y)/i! = f(y)+\sum_{i=0}^{k_1} s_i(x)y^i+\sum_{i=1}^{k_2} r_{2i}(x)p(y)y^i. \tag{4}$$

Let Ass be a free associative algebra over R with free generators $\{y, x_1, x_2, \ldots\}$. Then (4) yields in Ass

$$\sum_{i=0}^{\infty} x_i f^{(i)}(y)/i! = f(y) + \sum_{i=0}^{k_1} s_i^{\sigma}(x) y^i + \sum_{i=1}^{k_2} r_{2i}^{\sigma}(x) p(y) y^i, \tag{5}$$

where σ is a linear map: $R[x] \to$ Ass,

$$g(x)^{\sigma} = \sum_{i=0}^{k} \alpha_i x_i, \quad \text{if} \quad g(x) = \sum_{i=0}^{k} \alpha_i x^i.$$

Finally, (5) implies

$$vf(t)a = vaf(t) + \sum_{i=0}^{k_1} [a, s_i(adt)] t^i + \sum_{i=1}^{k_2} [a, r_{2i}(adt)] p(t) t^i \in W_0,$$

since $[a, r_{2i}(adt)] \in L_0$, $[L, s(adt)] = 0$ and $s_i(x) = s(x) l_i(x)$. We constructed an L-module $U = W/W_0$, which is free as R-module since $f(x)$ is unitary polynomial. The L-module U is triangulable since all the roots of $f(x)$ are in R. ∎

With this lemma we can complete the inductive step and prove theorem in the case of L splitting.

Suppose now that L is not splitting. Recall that Q is the field of fraction of R. A Lie Q-algebra $L_Q = L \otimes_R Q$ is embedded to a splitting Q-algebra $\mathcal{L} = \mathcal{T} \oplus \mathcal{N}$. Here, $\mathcal{N} = \sum_{\alpha \in \tau} \oplus \mathcal{N}_{\alpha}$, where $\tau \subset Q^*$ and $\mathcal{N}_{\alpha} = \{n \in \mathcal{N} \mid nt = \alpha(t)n, \forall t \in \mathcal{T}\}$ (see [4]). Let $\{g_1, \ldots, g_n\} \subseteq L$ be free basis of L. In $\mathcal{L}$, we have $g_i = t_i + n_i$, $t_i \in \mathcal{T}$, $n_i \in \mathcal{N}$, $i = 1, \ldots, n$. Note that $g_1, \ldots, g_n \in T = \{t \in \mathcal{T} \mid \alpha(t) \in R, \forall \alpha \in \tau\}$ and T is a free R-module. In $\mathcal{L}$, we have $n_i = \sum_{\alpha \in \tau} n_{i\alpha}$, $n_{i\alpha} \in \mathcal{N}_{\alpha}$, $i = 1, \ldots, n$. Let N_1 be an R-algebra Lie with generators $\{n_{j\alpha} \mid \alpha \in \tau, i = 1, \ldots, n\}$. It is obvious that N_1 is a finite-dimensional R-module and $N_1 T \subset N_1$. Hence $L_1 = T \oplus N_1$ is a splitting completely solvable R-algebra and L is subalgebra of L_1. ∎

References

1. A. Grishkov, *Analytic alternative loops*. Preprint, N510, 1985, Novosibirsk.
2. ———, *Lie's theorem for soluble binary Lie Algebras*. Algebra and Logic (3) **20**(1981), 255–281.
3. N. Jacobson, *Lie Algebras*. Interscience Publishers, New York, 1962.
4. A. Malcev, *On solvable Lie algebras*. Izvest. Akad. Nauk S.S.R. Ser. Mat. **9**(1945), 329–356. Amer. Math. Soc. Transl. (27) **1950**(1950).
5. ———, *Analytic loops*. J. Mat. Sb. **36**(1955), 569–576.

Omsk State University (Russia) and University of São Paulo (Brazil)
email: grishkov@ime.usp.br

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 82–85

THE CENTRAL CRITICAL VALUE OF AUTOMORPHIC L-FUNCTIONS

J. SENGUPTA

Presented by M. Ram Murty, FRSC

ABSTRACT. In this note we study the growth of the central critical value of L-functions of holomorphic cusp forms of varying weight k. We prove that, on the average, the central critical value obeys the Lindelöf hypothesis in the k aspect. Under the assumption of the Lindelöf hypothesis, we also give a lower bound for the number of cusp forms that have nonvanishing central critical value.

RÉSUMÉ. Dans cette Note, nous étudions la croissance de la valuer critique centrale des fonctions L des formes modulaires paraboliques de poids variable k. Nous démontrons que cette valuer obéit, en moyenne, l'hypothèse de Lindelöf d'aspect k. Sous l'hypothèse de Lindelöf, nous donnons aussi une minoration du nombre des formes modulaires paraboliques dont la valuer critique centrale ne s'annule pas.

The central critical value of L-functions for GL(2) cusp forms has been the subject of intensive research. It has various aspects like algebraicity, positivity, behaviour under twists, *etc.* In this note we will look at the L-series of holomorphic cusp forms of even integral weight. Most of the previous studies on this topic have been concerned with the cases when the weight is kept fixed and the level of the form is varied. This has been done in particular for forms of weight 2 and varying level in view of their connection with elliptic curves and the Birch-Swinnerton-Dyer conjecture (see [2], [6]).

In contrast to the above we will investigate the variation in the central critical values of L-functions of holomorphic cusp forms of varying weight and fixed level. In order to keep the exposition simple we will consider only forms of level one, *i.e.*, forms for the full modular group. Our approach is direct and our proofs are simple. In particular we do not make use of any approximate functional equations, trace formula for Hecke operators, *etc.*

NOTATION. Let k be an even integer ≥ 12 and let S_k denote the space of cusp forms of weight k for the full modular group. We denote by $\langle\, , \rangle$ the Petersson inner product in S_k and by g_k the dimension of S_k. Let $[f_\nu \mid 1 \leq \nu \leq g_k]$ denote the orthogonal basis of S_k consisting of normalised Hecke eigenforms and

Received by the editors June 20, 1999.
AMS subject classification: 11F11, 11F67.

let $L(f_\nu, s)$ denote the L-series of f_ν. It is well-known (see [5]) that $L(f_\nu, k/2)$ is non-negative. In the following we will assume that k is divisible by 4 for otherwise the central critical value is zero.

THEOREM. *For all* $\epsilon > 0$,

$$\sum_{\nu=1}^{g_k} L(f_\nu, k/2) \ll k^{1+\epsilon}.$$

COROLLARY. *On the average the central critical value obeys the Lindelöf hypothesis in the k aspect.*

PROOF OF THE THEOREM. Our proof is direct and is based on the explicit representation of a related sum given in [4] and the results of Hoffstein-Lockhart [3]. We begin by quoting the relevant result from [4].

PROPOSITION [4, p. 188]. *Let* $s \in \mathbb{C}$ *and let* $\sigma = \Re s$ *satisfy* $1 < \sigma < k-1$. *Then we have*

$$\begin{aligned}(2\pi)^s\Gamma(k-s) &+ (2\pi)^{k-s}\Gamma(s)\\ &+\frac{1}{2}(2\pi)^k\Bigg[\sum_{\substack{(a,c)\in\mathbb{Z}^2, ac>0\\ \gcd(a,c)=1}} c^{-k}\left(\frac{c}{a}\right)^s\\ &\quad\cdot\left(e^{2\pi i a'/c}e_1^{\pi i s/2}f_1(s,k;-2\pi i/ac) + e^{-2\pi i a'/c}e_1^{-\pi i s/2}f_1(s,k;2\pi i/ac)\right)\Bigg]\\ &= (2\pi)^{-\bar{s}}\Gamma(\bar{s})c_k\sum_{\nu=1}^{g_k}\frac{L(f_\nu,\bar{s})}{\langle f_\nu, f_\nu\rangle}.\end{aligned}$$

Here a' is the inverse of a modulo c, ${}_1f_1(\alpha,\beta;z) = \frac{\Gamma(\beta-\alpha)\Gamma(\alpha)}{\Gamma(\beta)}\,{}_1F_1(\alpha,\beta;z)$ and ${}_1F_1$ is Kummer's confluent hypergeometric function and

$$c_k := \frac{\pi\Gamma(k-1)}{2^{k-2}}.$$

Evaluating the above equality for $s = \frac{k}{2}$ we obtain

$$\begin{aligned}\sum_{\nu=1}^{g_k}&\frac{L(f_\nu,k/2)}{\langle f_\nu, f_\nu\rangle}\\ &= \frac{(2\pi)^{k/2}}{\Gamma(k/2)c_k}\Bigg\{(2\pi)^{k/2}\Gamma(k/2) + (2\pi)^{k/2}\Gamma(k/2) + \frac{1}{2}(2\pi)^k\\ &\qquad\times\sum_{\substack{(a,c)\in\mathbb{Z}^2, ac>0\\ \gcd(a,c)=1}} c^{-k}\left(\frac{c}{a}\right)^{k/2}\\ &\qquad\cdot\left(e^{2\pi i a'/c}e^{\pi i k/4}\,{}_1f_1(k/2,k;-2\pi i/ac)\right.\\ &\qquad\left.+e^{-2\pi i a'/c}e^{-\pi i k/4}\,{}_1f_1(k/2,k;2\pi i/ac)\right)\Bigg\}.\end{aligned}$$

Now ${}_1f_1(k/2, k; \pm 2\pi i/ac) = \int_0^1 e^{\pm 2\pi i u/ac} u^{k/2-1}(1-u)^{k/2-1}\, du$.

Hence $|{}_1f_1(k/2, k; \pm 2\pi i/ac)| \le B(k/2, k/2) = \frac{\Gamma^2(k/2)}{\Gamma(k)}$.

Using the above estimate and Legendre's duplication formula for the Γ function we obtain,

$$\sum_{\nu=1}^{g_k} \frac{L(f_\nu, k/2)}{\langle f_\nu, f_\nu\rangle} \ll \frac{(2\pi)^k}{c_k}$$

where the implied constant is absolute. Now

$$\|f_\nu\|^2 = \frac{\Gamma(k)}{2^{2k-1}\pi^{k+1}} L(1, F_\nu).$$

Here F_ν is the "adjoint square lift" of f_ν and $L(s, F_\nu)$ its standard L-function.

The result of Hoffstein-Lockhart ([3] and appendix of [3]) shows that $\forall \epsilon > 0$,

$$\frac{\Gamma(k)}{2^{2k-1}\pi^{k+1}} \frac{c_1}{\log(k+1)} \le \|f_\nu\|^2 \ll \frac{\Gamma(k)k^\epsilon}{2^{2k-1}\pi^{k+1}}$$

where c_1 is an effective positive constant. Substituting the expression for c_k and using the estimate above for the Petersson norm of f_ν we get

$$\sum_{\nu=1}^{g_k} L(f_\nu, k/2) \ll k^{1+\epsilon}. \qquad \blacksquare$$

REMARK. The implied constant in our theorem is effective.

Our next result deals with the number of cusp forms that have non-vanishing, hence positive, central critical value. In order to derive it we will assume the Lindelöf hypothesis in the k aspect for the individual f_ν's, *i.e.*, $L(f_\nu, k/2) \ll k^\epsilon$ $\forall \epsilon > 0$.

PROPOSITION. *Let* $0 < \epsilon < 1$ *be fixed but arbitrary. There exists a positive constant* C_ϵ*, depending only on* ϵ*, such that for all* k *sufficiently large, at least* $C_\epsilon k^{1-\epsilon}$ *of the* f_ν*'s have positive central critical value.*

PROOF OF THE PROPOSITION. For brevity we denote $L(f_\nu, k/2)$ by α_ν and $\|f_\nu\|^{-2}$ by p_ν.

Our assumption then is $\alpha_\nu \ll k^\epsilon$. The proof of the theorem above shows that

$$\sum_{\nu=1}^{g_k} p_\nu \alpha_\nu \ge 2\frac{(2\pi)^k}{c_k}\left(1 - \frac{(2\pi)^{k/2}\zeta^2(k/2)\Gamma(k/2)}{\zeta(k)\Gamma(k)}\right).$$

Using Legendre's duplication formula for $\Gamma(k) = \Gamma(2 \times k/2)$ we see that, for all k sufficiently large,

$$\sum_{\nu=1}^{g_k} p_\nu \alpha_\nu \ge \frac{(2\pi)^k}{c_k}.$$

On the other hand we have by [3] $p_\nu \le \frac{2^{2k-1}\pi^{k+1}}{c_1(k-1)\Gamma(k-1)} \log(k+1)$.

These two facts immediately lead to our proposition.

REMARK. If we assume that the constant in the Lindelöf hypothesis is effective, then the constant in our proposition is also effective.

REFERENCES

1. N. Abramowitz and I. Stegun, *Handbook of mathematical functions*. Dover, New York, 1965.
2. W. Duke, *The critical order of vanishing of automorphic L-functions with large level*. Invent. Math. **119**(1995), 165–174.
3. J. Hoffstein and P. Lockhart, *Coefficients of Maass forms and the Siegel zero*. Ann. of Math. **140**(1994), 161–181; also the appendix of this paper by D. Goldfeld, J. Hoffstein and D. Lieman.
4. W. Kohnen, *Nonvanishing of Hecke L-functions associated to cusp forms inside the critical strip*. J. Number Theory **67**(1997), 182–189.
5. W. Kohnen and D. Zagier, *Values of L-series of modular forms at the centre of the critical strip*. Invent. Math. **64**(1981), 175–198.
6. J. VanderKam, *Non-vanishing of L functions at the center of the critical strip*. Thesis, Princeton, 1998.

School of Mathematics
Tata Institute of Fundamental Research
Bombay 400005
India
email: sengupta@math.tifr.res.in

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 86–91

COUNTEREXAMPLES TO THE THEOREM OF PARKER

OLEG I. BOGOYAVLENSKIJ

Presented by Vlastimil Dlab, FRSC

ABSTRACT. Exact global helically symmetric solutions to the plasma equilibrium equations are derived. The solutions model astrophysical jets and provide counterexamples to a well-known theorem of Parker.

RÉSUMÉ. Des solutions globales exactes héliquement symmétriques aux équations d'équilibre de plasma sont construites. Ces solutions peuvent modeler des jets astrophysiques, et elles fournissent des contre-exemples à une théorème bien connue de Parker.

The problem of finding the nontrivial global solutions to the plasma equilibrium equations

$$\mathbf{J} \times \mathbf{B} = \operatorname{grad} p, \quad \operatorname{div} \mathbf{B} = 0, \quad \mathbf{J} = \operatorname{curl} \mathbf{B}/\mu \tag{1}$$

(where $\mathbf{B}$ is the magnetic field, $\mathbf{J}$ is the electric current density, p is the plasma pressure, and μ is the magnetic permeability) is known since their first applications in the 1950's to the problem of controlled thermonuclear fusion [5], [7] and to the astrophysical problems. However, all found until now exact solutions to equations (1) which are not translationally invariant either have singularities or are not localized and unboundedly grow at infinity [1], [3], [6]. The physically relevant global plasma equilibria have to satisfy the following conditions in the cylindrical coordinates r, ϕ, z: a) The magnetic field $\mathbf{B}$ and pressure p are smooth and bounded in $\mathbb{R}^3$; b) At $r \longrightarrow \infty$, the magnetic field $\mathbf{B} \longrightarrow 0$, the pressure $p \longrightarrow p_1$; c) All magnetic field lines are bounded in the radial variable r.

The plasma equilibrium equations (1) have two important reductions derived by Grad and Shafranov for the axially symmetrics case and by Johnson, Frieman, Kulsrud and Oberman (JFKO) [5] for the helically symmetric case. In the present paper, we report some new exact solutions to the JFKO equation which have astrophysical and laboratory applications.

We derive solutions to the JFKO equation [5] which are defined in the whole Euclidean space $\mathbb{R}^3$ and have no singularities. These exact plasma equilibria along with the exact axially symmetric solutions [2] model the variety of magnetic fields in astrophysical jets [4] and in solar prominences. The asymptotic value

Received by the editors August 20, 1999.
AMS subject classification: 35C05, 85A30.

of pressure p_1 in condition b) is the average pressure in the external medium (the astrophysical outflow or solar coronal plasma). The plasma equilibria are localized in the sense that the total magnetic energy in any layer $c_1 < z < c_2$ is finite. The magnetic field decreases at $r \longrightarrow \infty$ as rapidly as $c_N \exp(-\beta r^2) r^{2N}$. Hence the high collimation of the jet follows.

The obtained exact equilibrium solutions provide a helically symmetric counterexample to the well-known Parker's theorem [9] that has been around for more than 25 years and that concerns the small perturbations of the z-invariant plasma equilibria. Parker writes [9, p. 374]: "Consider a magnetic field $B_i(x,y) + \epsilon b_i(x,y,z)$ in the neighbourhood of the general equilibrium field $B_i(x,y)$"; and after a detailed study arrives at the conclusion on p. 377:

> Thus, in the general case, we are led to the conclusion that the invariance $\partial b_i/\partial z = 0$ (14.51) is a necessary condition for equilibrium. Any field in which winding pattern changes along the field, so that (14.51) is excluded by the topology, cannot be in equilibrium.

In [8], [11], this conclusion was called *Parker's theorem*. Many consequences and generalizations were produced assuming that the theorem is true, see [8], [10], [11], [12] and references therein.

We present a counterexample to Parker's theorem which satisfies all Parker's conditions [9, pp. 359–391], and the above physical conditions a), b), c). We construct a family of global z-invariant plasma equilibria, each equilibrium possesses a 3-dimensional linear space of helically symmetric global perturbations. The most important feature of these exact solutions is that they do *depend* on variable z and hence they are *not* z-invariant.

The helically symmetric magnetic field $\mathbf{B}$ has the form $\mathbf{B} = (\psi_u \hat{\mathbf{e}}_r - \psi_r \hat{\mathbf{e}}_u + f\hat{\mathbf{e}}_\phi)/r$, where $\psi = \psi(r,u)$, $f = f(r,u)$, $\psi_x = \partial\psi/\partial x$, and $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_u$, $\hat{\mathbf{e}}_\phi$ are the unit tangent vectors corresponding to the coordinates $r, u = z - \gamma\phi, \phi$. Johnson *et al.* had shown in [5] that the plasma equilibrium equations (1) for the helically symmetric solutions are equivalent to the equalities $(r^2+\gamma^2)f/r^2 - \gamma\psi_r/r = I(\psi)$, $p = p(\psi)$ with arbitrary functions $I(\psi)$ and $p(\psi)$, and the JFKO equation

$$\frac{1}{r^2}\frac{\partial^2\psi}{\partial u^2} + \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r}{F}\frac{\partial\psi}{\partial r}\right) + \frac{II'(\psi)}{F} + \frac{2\gamma I(\psi)}{F^2} = -\mu p'(\psi), \tag{2}$$

where prime denotes differentiation with respect to ψ and $F = r^2 + \gamma^2$. The magnetic field $\mathbf{B}$ has the following form in the cylindrical coordinates r, z, ϕ:

$$\mathbf{B} = \psi_u \hat{\mathbf{e}}_r/r + \big(\gamma I(\psi) - r\psi_r\big)\hat{\mathbf{e}}_z/F + \big(rI(\psi) + \gamma\psi_r\big)\hat{\mathbf{e}}_\phi/F. \tag{3}$$

THEOREM 1. *There exist an infinite family of global plasma equilibria depending on an arbitrary integer $N \geq 0$ and two reals $\beta > 0$ and γ. The equilibria are z- and ϕ-invariant in the cylindrical coordinates r, z, ϕ. For any two integers $m \geq 1$ and $n \geq 0$ satisfying the condition $2n + m < N$ and parameter*

$\beta = m^2/(4\gamma^2(2N - 2n - m))$, *the equilibria are contained in a 3-dimensional linear family of global plasma equilibria which are invariant with respect to the helical transformations* $z \longrightarrow z + \gamma\phi_0$, $\phi \longrightarrow \phi + \phi_0$.

PROOF. Suppose that $I(\psi) = \alpha\psi$ and $p(\psi) = p_1 - 2\beta^2\psi^2/\mu > 0$, where $p_1 > (2\beta^2/\mu)\max(\psi^2(x, y, z))$. Under these assumptions, the JFKO equation (2) becomes linear. Separating variables by the substitution $\psi(r, u) = A(r)(a\cos(\omega u) + b\sin(\omega u))$, we obtain the equation

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r}{F}\frac{\mathrm{d}A(r)}{\mathrm{d}r}\right) = \left(4\beta^2 + \frac{\omega^2}{r^2} - \frac{\alpha^2}{F} - \frac{2\alpha\gamma}{F^2}\right)A(r). \tag{4}$$

To find exact solutions, we substitute $A(r) = r^m e^{-\beta r^2}B(x)$, $x = 2\beta r^2$. For the non-singular solutions, we necessarily have $\gamma\omega = m$, where $m \geq 0$ is an integer. Suppose that

$$\alpha^2\gamma^2 = (m + c_1)^2 + 4nc_1, \tag{5}$$

where n is a parameter, $c_1 = 2\beta\gamma^2$. Then equation (4) turns into the equation

$$\begin{aligned}(x^2 + c_1 x)B'' + (-x^2 + (m - c_1)x + (m + 1)c_1)B' \\ + n(x + c_1 - k_{mn}c_1)B = 0,\end{aligned} \tag{6}$$

where $k_{mn} = (m + c_1 - \alpha\gamma)/(2nc_1)$. Equation (6) has a polynomial solution $B(x)$ only if parameter n is an integer $n \geq 0$. Substituting $B(x) = P(x) - k_{mn}xP'(x)$, we reduce equation (6) to $xP'' + (1 + m - x)P' + nP = 0$. This equation has polynomial solutions $P_{mn}(x) = \mathrm{d}^m L_{m+n}(x)/\mathrm{d}x^m$ of degree n where $L_p(x)$ are the classical Laguerre polynomials $L_p(x) = e^x \mathrm{d}^p(e^{-x}x^p)/\mathrm{d}x^p/p!$. Hence we obtain that if the two integers m and n satisfy equation (5) then equation (6) has the polynomial solution $B_{mn}(x) = \mathrm{d}^m L_{m+n}(x)/\mathrm{d}x^m - k_{mn}x\mathrm{d}^{m+1}L_{m+n}(x)/\mathrm{d}x^{m+1}$. Therefore the JFKO equation (2) has the exact solutions

$$\psi_{mn} = r^m e^{-\beta r^2}B_{mn}(2\beta r^2)(a_{mn}\cos(mu/\gamma) + b_{mn}\sin(mu/\gamma)). \tag{7}$$

For $m \geq 2$, the magnetic field (3) $\mathbf{B}_{mn}(0, z) = 0$ on the axis $r = 0$. Only the $m = 1$ kink mode (7) has nonzero magnetic field $\mathbf{B}_{mn}(0, z)$ at $r = 0$.

For $m = n = 0$, equation (5) implies $\alpha = 2\beta\gamma$, and equation (6) has solution $B = \text{const}$. Hence equation (2) has an exact solution of the Gaussian form $\psi_0(r) = \exp(-\beta r^2)$. For $m = 0$ and $n = N \neq 0$, equation (5) becomes $\alpha^2\gamma^2 = c_1^2 + 4Nc_1$ and the exact solution (7) takes the form

$$\psi_N(r) = a_N e^{-\beta r^2}B_{0N}(2\beta r^2), \tag{8}$$

where polynomials $B_{0N}(x)$ are $B_{0N}(x) = L_N(x) - k_N x L'_N(x)$, and $k_N = (1 - \sqrt{1 + 2N/(\beta\gamma^2)})/(2N) < 0$. These formulae and the known properties of the Laguerre polynomials $L_N(x)$ imply that each polynomial $B_{0N}(x)$ has N distinct positive roots. Hence all its roots are simple.

For the flux function (8), the magnetic field (3) has no singularities and decreases at $r \longrightarrow \infty$ as rapidly as $\exp(-\beta r^2)r^{2N}$. The magnetic sufraces $\psi_N(r) =$ const are cylinders $r =$ const. Thus equation (2) has global solutions (8) where the two parameters $\beta > 0$ and $\gamma \neq 0$ are arbitrary. These plasma equilibria possess cylindrical symmetry for they are z- and ϕ-invariant.

EXACT GLOBAL PLASMA EQUILIBRIA. For $m \geq 1$, the flux functions $\psi_{mn}(r, u)$ (7) define the magnetic fields $\mathbf{B}_{mn}$ (3) that are smooth in the cartesian coordinates and decrease at $r \longrightarrow \infty$ as rapidly as $c(u)\exp(-\beta r^2)r^{2n+m}$. Hence these solutions satisfy the above conditions a) and b).

To find the global equilibria, we consider a linear combination of the exact solutions $\psi_N(r)$ (8) and $\psi_{mn}(r, u)$ (7) which satisfies equation (2) if equations $\alpha^2\gamma^2 = c_1^2 + 4Nc_1$ and (5) hold simultaneously. These two equations yield the inequality $2N > 2n + m$ and the formulae $c_1 = 2\beta\gamma^2 = m^2/q$, $|\alpha\gamma| = m\sqrt{(4N-m)^2 - 16nN}/q$, where $q = 2(2N - 2n - m)$. Thus we obtain the exact solutions to the JFKO equation (2):

$$(9) \quad \psi_{Nmn} = e^{-\beta r^2}\Big(a_N B_{0N} + r^m B_{mn}\big(a_{mn}\cos(mu/\gamma) + b_{mn}\sin(mu/\gamma)\big)\Big),$$

where N, m, n are arbitrary integers ≥ 0 satisfying the inequality $2N > 2n + m$. This inequality implies that function $\psi_{Nmn}(r, u)$ (9) has the leading term $(-2\beta)^N a_N \exp(-\beta r^2)r^{2N}$ at $r \longrightarrow \infty$. Hence equilibria (9) for $m \geq 1$ satisfy the above conditions a) and b), and all magnetic surfaces $\psi_{Nmn}(r, u) =$ const asymptotically for $r >> 1$ are cylinders $r =$ const. Therefore all magnetic field lines are bounded in variable r. Thus the exact solutions (9) define the global plasma equilibria. For $m \geq 1$, these exact solutions describe the highly collimated astrophysical jets. ∎

In the coordinates r, z, ϕ, the magnetic surfaces $\psi(r, u) = h =$ const are the helically rotating cylinders C_h. The z-periodicity of $\psi_{Nmn}(r, u)$ implies that the sections S_k: $z = 2\pi\gamma k$ of a cylinder C_h are the same for all integers k. Hence the subsequent intersections of magnetic field lines on C_h with curves S_0 and S_1 define a rotation tranform $T_h: S_0 \longrightarrow S_0 = S_1$. For an appropriate angular parameter $s = s \bmod 1$ on the closed curve S_0, the mapping T_h is $T_h(s) = s + \theta_0(h)$. For a generic h, the rotation number $\theta_0(h)$ is irrational. Hence the generic magnetic field lines are quasi-periodic in variables r, z, ϕ which implies that the magnetic field lines never repeat in the z-direction, but can have a structure arbitrarily close to the initial data. Hence their "winding pattern" changes continuously with z, and does not repeat.

COUNTEREXAMPLES TO PARKER'S THEOREM. The exact solution (8) defines the magnetic field (3): $\mathbf{B}_N(r) = \big((\alpha\gamma\psi_N - r\psi'_N)\hat{\mathbf{e}}_z + (\alpha r\psi_N + \gamma\psi'_N)\hat{\mathbf{e}}_\phi\big)/(r^2 + \gamma^2)$, that is nonzero everywhere in the Euclidean space $\mathbb{R}^3$. Indeed, an equality $\mathbf{B}_N(r_0) = 0$ implies $\psi_N(r_0) = \psi'_N(r_0) = 0$, hence $B_{0N}(x_0) = B'_{0N}(x_0) = 0$, a contradiction, because all roots of the polynomial $B_{0N}(x)$ are simple. Hence for any R and all $r \leq R$, we have $|\mathbf{B}_N(r)| \geq a_N B(R) > 0$.

Let $\mathbf{B}_{Nmn}(r,u)$ be the magnetic field defined by the flux function ψ_{Nmn} (9) and $A_N = (|a_{mn}| + |b_{mn}|)/|a_N|$. Inside any domain $0 \leq r \leq R$, we have

$$|\mathbf{B}_{Nmn} - \mathbf{B}_N|/|\mathbf{B}_N| < A_N C(R)/B(R). \tag{10}$$

At $r \longrightarrow \infty$, the asymptotics hold:

$$\frac{|\psi_{Nmn} - \psi_N|}{|\psi_N|} < A_N(2\beta)^{n-N}\frac{2C_{Nmn}}{r^{2N-2n-m}}, \quad C_{Nmn} = \frac{(1-nk_{mn})N!}{(1-Nk_N)n!}. \tag{11}$$

The inequality $2N > 2n+m$ implies $|\psi_{Nmn} - \psi_N|/|\psi_N| \longrightarrow 0$ at $r \longrightarrow \infty$. Hence for $A_N << 1$, we obtain $|\mathbf{B}_{Nmn} - \mathbf{B}_N|/|\mathbf{B}_N| << 1$ everywhere in $\mathbb{R}^3$.

Let x_N be the greatest root of the polynomial $B_{0N}(x)$. For $x > x_N$, we have $|B_{0N}(x)| > (1 - Nk_N)(x - x_N)^N/N!$ and $|B_{mn}(x)| < (1 - nk_{mn})x^n/n!$. Hence we get $|\psi_{Nmn} - \psi_N|/|\psi_N| < A_N C_{Nmn} r^m x^n/(x - x_N)^N$. Hence for $x > N^{2\kappa+1}x_N$, $2\kappa = |2\log C_{Nmn} - m\log(2\beta)|/(2N-2n-m)\log N$, we obtain $|\psi_{Nmn} - \psi_N|/|\psi_N| < A_N$. The same inequality is true for the magnetic field. Thus for $A_N << 1$, the perturbations (9) can be significant only for $x < N^{2\kappa+1}x_N$. Substituting $x = 2\beta\ell^2$, we find for the characteristic length scale ℓ in the variable r of the perturbations (9): $\ell \leq N^\kappa\sqrt{Nx_N/2\beta}$.

The inequalities (10) and (11) mean that the plasma equilibria (9) at $A_N << 1$ are small perturbations in the whole Euclidean space $\mathbb{R}^3$ of the z-invariant equilibrium (8). Hence we obtain that Parker's condition that "the local perturbation to the field is small compared to the total field" [9, p. 361] is satisfied everywhere. Parker's condition that "the magnetic field is analytic in its deviation ϵ from the invariant field $B_i(x,y)$" [9, p. 378] is satisfied because the exact solutions (9) are linear functions of the small parameters a_{mn}, b_{mn}. Parker's condition that the length of the flux tube L is "large compared to the characteristic transverse scale of variation ℓ of the field" [9, p. 362] is satisfied because $\ell \leq N^\kappa\sqrt{Nx_N/2\beta}$ and the flux tube length L can be taken arbitrarily large for the z-invariant equilibrium (9). Hence $L >> \ell$. All perturbations (9) are *not* z-invariant. Hence the plasma equilibria (9) are counterexamples to Parker's theorem.

ACKNOWLEDGEMENT. The author thanks R. M. Erdahl and R. N. Henriksen for helpful discussions.

REFERENCES

[1] D. Biskamp, *Nonlinear magnetohydrodynamics*. Cambridge University Press, Cambridge, 1993.

[2] O. I. Bogoyavlenskij, *Astrophysical jets as exact plasma equilibria*. Phys. Rev. Lett. **84**(2000), 1914–1917.

[3] P. K. Browning, *Magnetohydrodynamics in solar coronal and laboratory plasmas: a comparative study*. Phys. Rep. **169**(1988), 329–384.

[4] D. C. Hines, F. N. Owen and J. A. Eilek, *Filaments in the radio lobes of M87.* Astrophys. J. **347**(1989), 713–726.

[5] J. L. Johnson, C. R. Oberman, R. M. Kulsrud and E. A. Frieman, *Some stable hydromagnetic equilibria.* Phys. Fluids **1**(1958), 281–296.

[6] B. B. Kadomtsev, *Equilibrium of a plasma with helical symmetry.* Soviet Phys. JETP, **10**(1960), 962–963.

[7] M. D. Kruskal and R. M. Kulsrud, *Equilibrium of a magnetically confined plasma in a toroid.* Phys. Fluids, **1**(1958), 265 274.

[8] B. C. Low, *Exact static equilibrium of vertically oriented magnetic flux tubes.* Solar Phys. **67**(1980), 57–77.

[9] E. N. Parker, *Cosmical magnetic fields.* Clarendon Press, Oxford, 1979.

[10] R. Rosner and E. Knobloch, *On perturbations of magnetic field configurations.* Astrophys. J. **262**(1982), 349–357.

[11] K. C. Tsinganos, *Magnetohydrodynamic equilibrium. IV. Nonequilibrium of nonsymmetric hydrodynamic topologies.* Astrophys. J. **259**(1982), 832–843.

[12] M. Villata and K. Tsinganos, *Exact solutions for helical magnetohydrodynamics equilibria.* Phys. Fluids B **5**(1993), 2153–2164.

Department of Mathematics and Statistics
Queen's University
Kingston, Ontario
K7L 3N6
email: Bogoyavl@oib.mast.Queens.ca

C. R. Math. Rep. Acad. Sci. Canada Vol. **22** (2), 2000 pp. 92–96

TRACES OF LOCALIZATION OPERATORS

JINGDE DU AND M. W. WONG

Presented by L. A. Lorch, FRSC

ABSTRACT. The trace of any integral power of a localization operator on a separable and complex Hilbert space is computed and used to estimate the norm of a positive localization operator in the Schatten-von Neumann class S_p when p is a positive integer.

RÉSUMÉ. La trace d'une puissance arbitraire d'un opérateur de localisation sur un espace de Hilbert complexe et séparable est calculée et utilisée pour estimer la norme d'un opérateur positif de localisation dans la classe Schatten-von Neumann S_p où p est un nombre entier positif.

1. Introduction. Let G be a locally compact and Hausdorff group on which the left Haar measure is denoted by μ. For $1 \leq p \leq \infty$, we denote the norm in $L^p(G)$ by $\| \ \|_p$. Let X be a separable and complex Hilbert space of which the dimension is infinite. We denote the inner product and norm in X by $(\, , \,)$ and $\| \ \|$ respectively. Let $B(X)$ be the C^*-algebra of all bounded linear operators on X, and let $\| \ \|_*$ denote the norm in $B(X)$. An irreducible and unitary representation $\pi: G \to B(X)$ is said to be square-integrable if there exists a nonzero element φ in X such that

$$\int_G |(\varphi, \pi(g)\varphi)|^2 \, d\mu(g) < \infty. \tag{1.1}$$

We call any element φ in X for which $\|\varphi\| = 1$ and (1.1) is valid an admissible wavelet for the square-integrable representation $\pi: G \to B(X)$, and we define the constant c_φ by

$$c_\varphi = \int_G |(\varphi, \pi(g)\varphi)|^2 \, d\mu(g).$$

Let φ be an admissible wavelet for the square-integrable representation $\pi: G \to B(X)$. Then it can be proved that the resolution of the identity formula, *i.e.*,

$$(x, y) = \frac{1}{c_\varphi} \int_G (x, \pi(g)\varphi)(\pi(g)\varphi, y) \, d\mu(g), \quad x, y \in X,$$

Received by the editors September 23, 1999.

This research has been partially supported by the Natural Sciences and Engineering Research Council of Canada under Grant OGP0008562.

AMS subject classification: 47G10.

is valid. The theory of square-integrable representations hitherto described can be found in Grossmann, Morlet and Paul [1], [2], Holschneider [4] and Wong [7], [8] among others.

Let F be a function in $L^1(G)$ or $L^\infty(G)$. Then, for any x in X, we define $L_F x$ by

$$(1.2)\qquad (L_F x, y) = \frac{1}{c_\varphi}\int_G F(g)\big(x, \pi(g)\varphi\big)\big(\pi(g)\varphi, y\big)\, d\mu(g), \quad y \in X.$$

Then we have the following result proved in the paper [3] by He and Wong.

THEOREM 1.1. *Let $F \in L^p(G)$, $1 \le p \le \infty$. Then there exists a unique bounded linear operator $L_F: X \to X$ such that*

$$\|L_F\|_* \le c_\varphi^{-\frac{1}{p}} \|F\|_p$$

and $L_F x$ is given by (1.2) for all x in X and all simple functions F on G satisfying

$$\mu\{g \in G : F(g) \ne 0\} < \infty.$$

Let A be a compact operator from X into X. If we denote by $A^*: X \to X$ the adjoint of $A: X \to X$, then the linear operator $(A^*A)^{\frac{1}{2}}: X \to X$ is positive and compact. Let $\{\psi_k : k = 1, 2, \ldots\}$ be an orthonormal basis for X consisting of eigenvectors of $(A^*A)^{\frac{1}{2}}: X \to X$, and let $s_k(A)$ be the eigenvalue corresponding to the eigenvector ψ_k, $k = 1, 2, \ldots$. We say that the compact operator $A: X \to X$ is in the Schatten-von Neumann class S_p, $1 \le p < \infty$, if

$$\sum_{k=1}^{\infty} s_k(A)^p < \infty.$$

It can be shown that S_p, $1 \le p < \infty$, is a Banach space in which the norm $\|\ \|_{S_p}$ is given by

$$\|A\|_{S_p} = \Big\{\sum_{k=1}^{\infty} s_k(A)^p\Big\}^{\frac{1}{p}}, \quad A \in S_p.$$

We let S_∞ be the C^*-algebra $B(X)$ of all bounded linear operators from X into X.

The following result is Theorem 6.1 in the paper [3] by He and Wong.

THEOREM 1.2. *Let $F \in L^p(G)$, $1 \le p \le \infty$. Then the bounded linear operator $L_F: X \to X$ is in S_p and*

$$\|L_F\|_{S_p} \le \left(\frac{4}{c_\varphi}\right)^{\frac{1}{p}} \|F\|_p.$$

In this paper, we compute the trace of a localization operator $L_F: X \to X$ and estimate the trace of any integral power of $L_F: X \to X$ for any function F in $L^1(G)$. As a supplement to Theorem 1.2, we give another estimate on the norm of the localization operator $L_F: X \to X$ in the Schatten-von Neumann class S_p when F is a nonnegative function in $L^1(G)$ and p is a positive integer.

The motivation for the trace formula given in this paper comes from the study of the free energy and the partition function in quantum statistical mechanics described in, *e.g.*, the paper [5] by Peierls, and the most interesting groups to which the theory of this paper applies are the Weyl-Heisenberg group and the affine group.

2. The trace. The starting point is the following proposition, which can be found on page 211 of the book [6] by Reed and Simon.

PROPOSITION 2.1. *Let $A: X \to X$ be a bounded linear operator in S_1 and let $\{\varphi_k : k = 1, 2, \ldots\}$ be any orthonormal basis for X. Then the series $\sum_{k=1}^{\infty}(A\varphi_k, \varphi_k)$ is absolutely convergent and the sum is independent of the choice of the orthonormal basis $\{\varphi_k : k = 1, 2, \ldots\}$.*

In view of Proposition 2.1, we can define the trace $\operatorname{tr}(A)$ of any linear operator $A: X \to X$ in S_1 by

$$\operatorname{tr}(A) = \sum_{k=1}^{\infty}(A\varphi_k, \varphi_k), \tag{2.1}$$

where $\{\varphi_k : k = 1, 2, \ldots\}$ is any orthonormal basis for X.

Let $F \in L^1(G)$. Then the localization operator $L_F: X \to X$ is in S_1 and we have the following interesting result on its trace.

THEOREM 2.2. $\operatorname{tr}(L_F) = \frac{1}{c_\varphi} \int_G F(g)\, d\mu(g)$.

PROOF. Let $\{\varphi_k : k = 1, 2, \ldots\}$ be an orthonormal basis for X. Then, using (1.2), (2.1), Fubini's theorem, the Parseval identity, $\|\varphi\| = 1$ and the fact that $\pi(g): X \to X$ is a unitary operator for all g in G, we get

$$\begin{aligned}
\operatorname{tr}(L_F) &= \sum_{k=1}^{\infty}(L_F\varphi_k, \varphi_k) \\
&= \sum_{k=1}^{\infty} \frac{1}{c_\varphi} \int_G F(g) \left|(\varphi_k, \pi(g)\varphi)\right|^2 d\mu(g) \\
&= \frac{1}{c_\varphi} \int_G F(g) \sum_{k=1}^{\infty} \left|(\varphi_k, \pi(g)\varphi)\right|^2 d\mu(g) \\
&= \frac{1}{c_\varphi} \int_G F(g) \|\pi(g)\varphi\|^2 \, d\mu(g) \\
&= \frac{1}{c_\varphi} \int_G F(g)\, d\mu(g).
\end{aligned}$$

∎

Since S_1 is an ideal in $B(X)$, it follows that, for $m = 1, 2, \ldots$, the linear operator $L_F^m: X \to X$ is in S_1. The following theorem gives an estimate for the trace $\operatorname{tr}(L_F^m)$ of $L_F^m: X \to X$ in terms of the trace $\operatorname{tr}(L_{|F|})$ of the localization operator $L_{|F|}: X \to X$.

THEOREM 2.3. *For* $m = 1, 2, \ldots$,

$$|\operatorname{tr}(L_F^m)| \leq \left(\operatorname{tr}(L_{|F|})\right)^m.$$

PROOF. Let $\{\varphi_k : k = 1, 2, \ldots\}$ be an orthonormal basis for X. Then, using (1.2), Theorem 1.1, (2.1), Fubini's theorem, the Schwarz inequality, the Parseval identity, $\|\varphi\| = 1$, $L_F^* = L_{\bar{F}}$ and the fact that $\pi(g): X \to X$ is a unitary operator for all g in G, we get

$$\begin{aligned}
|\operatorname{tr}(L_F^m)| &= \left|\sum_{k=1}^{\infty}(L_F^m\varphi_k, \varphi_k)\right| = \left|\sum_{k=1}^{\infty}(L_F\varphi_k, L_{\bar{F}}^{m-1}\varphi_k)\right| \\
&= \left|\sum_{k=1}^{\infty}\frac{1}{c_\varphi}\int_G F(g)(\varphi_k, \pi(g)\varphi)(\pi(g)\varphi, L_{\bar{F}}^{m-1}\varphi_k)\, d\mu(g)\right| \\
&= \left|\sum_{k=1}^{\infty}\frac{1}{c_\varphi}\int_G F(g)(\varphi_k, \pi(g)\varphi)(L_F^{m-1}\pi(g)\varphi, \varphi_k)\, d\mu(g)\right| \\
&\leq \frac{1}{c_\varphi}\int_G |F(g)|\Big(\sum_{k=1}^{\infty}|(\varphi_k, \pi(g)\varphi)|^2\Big)^{\frac{1}{2}}\Big(\sum_{k=1}^{\infty}|(L_F^{m-1}\pi(g)\varphi, \varphi_k)|^2\Big)^{\frac{1}{2}}\, d\mu(g) \\
&= \frac{1}{c_\varphi}\int_G |F(g)|\, \|\pi(g)\varphi\|\, \|L_F^{m-1}\pi(g)\varphi\|\, d\mu(g) \\
&\leq \left(\frac{1}{c_\varphi}\int_G |F(g)|\, d\mu(g)\right)^m. \qquad (2.2)
\end{aligned}$$

Thus, by Theorem 2.2 and (2.2), the proof is complete. ∎

3. The Schatten-von Neumann property. A consequence of Theorem 2.3 is the following result supplementing Theorem 1.2.

THEOREM 3.1. *Let F be a nonnegative function in $L^1(G)$. Then, for $p = 1, 2, \ldots$,*

$$\|L_F\|_{S_p} \leq \operatorname{tr}(L_F).$$

Moreover,

$$\|L_F\|_{S_1} = \operatorname{tr}(L_F).$$

PROOF. Since F is a nonnegative function in $L^1(G)$, it follows from (1.2) that $L_F: X \to X$ is a positive operator. Thus, the singular values of $L_F: X \to X$ coincide with the eigenvalues of $L_F: X \to X$. So,

$$\|L_F\|_{S_p} = \Big(\sum_{k=1}^{\infty}(L_F^p\psi_k, \psi_k)\Big)^{\frac{1}{p}}, \qquad (3.1)$$

where $\{\psi_k : k = 1, 2, \ldots\}$ is an orthonormal basis for X consisting of eigenvectors of $L_F : X \to X$. By (2.1), (3.1) and Theorem 2.3,

$$\|L_F\|_{S_p} = \left(\mathrm{tr}(L_F^p)\right)^{\frac{1}{p}} \leq \mathrm{tr}(L_F)$$

for $p = 1, 2, \ldots$. If $p = 1$, then, by (2.1) and (3.1),

$$\|L_F\|_{S_1} = \mathrm{tr}(L_F). \qquad \blacksquare$$

References

1. A. Grossmann, J. Morlet and T. Paul, *Transforms associated to square integrable group representations I: General results.* J. Math. Phys. **26**(1985), 2473–2479.
2. A. Grossmann, J. Morlet and T. Paul, *Transforms associated to square integrable group representations II: Examples.* Ann. Inst. Henri Poincaré **45**(1986), 293–309.
3. Z. He and M. W. Wong, *Localization operators associated to square integrable group representations.* Panamer. Math. J. (1) **6**(1996), 93–104.
4. M. Holschneider, *Wavelets: An Analysis Tool.* Oxford University Press, 1995.
5. R. Peierls, *On a minimum property of the free energy.* Phys. Rev. **54**(1938), 918–919.
6. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis, Revised and Enlarged Edition.* Academic Press, 1980.
7. M. W. Wong, *Weyl Transforms.* Springer-Verlag, 1998.
8. ———, *Localization Operators.* Lecture Notes, Seoul National University, 1999.

Department of Mathematics and Statistics
York University
4700 Keele Street
Toronto, Ontario
M3J 1P3
email: jingde@pascal.math.yorku.ca
mwwong@pascal.math.yorku.ca